

TRICKLET CONFERENCE

Model building in empirical translation studies

RWTH Aachen University

May 19/20, 2022



DFG-funded translation research project –
Project no. NE 1822/2-2

Table of contents

Keynotes

A relevance-theoretic account towards an empirical theory of translation <i>Fabio Alves</i>	6
Representation, Metarepresentation and Relevance: An Enactive Reassessment <i>Michael Carl</i>	7
Disciplinary integration and the critical role of sociocognition in theorizing translation <i>Sandra L. Halverson</i>	9
Epistemic virtues in (empirical) translation studies <i>Haidee Kotze</i>	13
Top-down and Bottom-Up Approaches to Modeling Translation Variation <i>Ekaterina Lapshinova-Koltunski</i>	15
The contribution of learner translation corpora to theory development in empirical translation studies: Overview and future prospects <i>Marie-Aude Lefer</i>	16
Aeons later <i>Moritz Schaeffer</i>	21
Is there still a role to play for corpora when translation theories turn cognitive? <i>Gert De Sutter</i>	22
Session 1	
Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters <i>Nannan Liu</i>	25
Embodied Metarepresentations <i>Nicolás Araneda Hinrichs, Maryam Foradi, Tariq Yousef, Susanne Triesch, Elisa Hartmann</i>	29
Session 2	
Building a communication-based analytical framework in situated translation training <i>Jun Yang</i>	32
Syntactic characterization of constrained language based on dependency treebanks of translated English and non-native English <i>Ruimin Ma, Yue Jiang</i>	34

Session 3	
Trados and the CRITT TPR-DB: Translation Process Research in an Ecologically Valid Environment <i>Longhui Zou, Michael Carl</i>	38
Note the Editor: A quantitative and qualitative study of editorial impact on literary translators' habitus and practice <i>Noa Marom Tsadok</i>	41
Literalization in the self-revision process of novice and experienced biomedical translators <i>Susana Valdez</i>	44
Using semi-structured interviews to revisiting the decision tree model for post-editing tasks: What can the language industry teach us? <i>Jean Nitzke, Carmen Canfora, Silvia Hansen-Schirra</i>	47
Session 4	
Explicitation and implicitation in L1><L2 student translation: A relevance-theoretic approach <i>Laura Aguiar</i>	49
Register Variation in Student Translations: A Study on Impersonal Verb Use in Multiple Translation Corpora <i>Tiffany Jandrain</i>	53
Metaphor awareness and metaphor translation competence: an empirical study on translation students <i>Charlène Meyers</i>	57
Session 5	
A new perspective on translation studies – translation as overcoming communication barriers <i>Sergio Hernández Garrido, Rebecca Schulz</i>	60
A syntactic dependency network approach to the study of translational language <i>Lu Fan</i>	63
Four explanatory variables in an empirical model of translation <i>Tatiana Serbina, Mario Bisiada, Stella Neumann</i>	68
Session 6	
Unique Item Hypothesis Revisited from the Perspective of Translation Directionality <i>Juhong Zhan, Yue Jiang, Jin Baicen, Tianyuan Yang</i>	72
Activity as a discriminant for direct and inverse translations: A corpus-based study of eight versions of Li Sao <i>Zuohao Xu, Yue Jiang</i>	77

Keynotes

KEYNOTE

A relevance-theoretic account towards an empirical theory of translation

Fabio Alves

Universidade Federal de Minas Gerais UFMG

In his application of Relevance Theory (Sperber & Wilson 1986/1995) to translation studies, Gutt (1991/2000) proposed an agenda in favour of a competence-oriented research of translation (CORT) to account for the human ability to express in one given language what had been expressed in another language. For Gutt, translation is a case of interlingual interpretive language use supported by the principle of relevance. Although, originally, Gutt's work was not empirically oriented, it paved the way for several empirically oriented attempts to model translation task-execution from a relevance-theoretic perspective. Such studies have focused, among other aspects, on the relationship between processing effort and cognitive effect (Alves & Gonçalves 2003, Alves 2007), on the impact of conceptual, procedural and hybrid encodings during the translation process (Alves & Gonçalves 2013) or on a framework for post-editing machine translation (Carl & Schaeffer 2019). Drawing on the above-mentioned works, this presentation aims at discussing how the relevance-theoretic framework can contribute to the formulation of an emerging model or theory of translation that is grounded in empirical evidence. It also addresses the issue of how (or if) an extended framework for an empirical theory of translation can be construed of which the relevance-theoretic account would be an integral part.

References

- Alves, F. (2007). Cognitive effort and contextual effect in translation: a relevance-theoretic approach. *Journal of Translation Studies*, 10, 18-35.
- Alves, F. & Gonçalves, J.L. (2003). A relevance theory approach to the investigation of inferential processes in translation. In Fabio Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 3–24). Amsterdam: John Benjamins.
- Alves, F. & Gonçalves, J.L (2013). Investigating the conceptual-procedural distinction in the translation process: A relevance-theoretic analysis of micro and macro translation units. *Target* 25(1), 107-124.
- Carl, M. & Schaeffer, M. (2019). Outline for a relevance theoretical model of machine translation post-editing. In Defeng Li, Victoria Lei & Yuanjiang He (Eds.), *Researching cognitive processes of translation, New Frontiers in Translation Studies Series* (pp. 49-67). Singapore: Springer.
- Gutt, E.-A. (1991/2000). *Translation and relevance. Cognition and context*. Oxford: Blackwell. (2nd edition, Manchester: St. Jerome).
- Sperber, D. & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Oxford: Blackwell.

KEYNOTE

Representation, Metarepresentation and Relevance: An Enactive Reassessment

Michael Carl
Kent State University

Around 40 years ago, Translation Process Research (TPR) set out to investigate "What goes on in the minds of translators" (Krings, 1986). Since then, several techniques have been deployed to illuminate the translator's black box, and numerous models have been suggested to explain the hidden translation processes. The Monitor Model (Schaffer and Carl 2013) stipulates that two concurrent (sets of) translation processes complement each other in the translator's mind. Automatic/horizontal processes are driven by priming mechanisms. They are modular and quick. Controlled/vertical processes are slower. They take into account a large(r) number of cognitive resources, including internal and external search and conscious/reflective thought.

In this talk I intend to re-assess the Monitor Model under an enactive view. Rather than seeing cognition as a manipulation of mental symbols, enactivism emphasizes the mutual constitution of the self and the environment through dynamic interaction, picturing this reciprocal adjustment in terms of largely non-representational cognitive processes. I will point to similarities between the Monitor Model and the Global Neuronal Workspace Theory (e.g., Dehaene 2014), one of the leading cognitive accounts of consciousness. Relevance Theory (RT) suggests a notion of relevance to be at the core of translational activity. According to RT, translation is a form of communication which follows the principle of relevance (Gutt 2000). The principle of relevance stipulates that participants in a communicative situation follow a path of least effort while aiming at producing maximum cognitive effects.

Supported by observable traces of translational effort and effects from translation process data, I will introduce a dynamic notion of translation relevance as a multivariate field of effort and effect indicators, in which the principle of relevance defines an optimal path. Conceptualizing relevance in terms of a multivariate field of effect and effort underpins the enactivist nature of translation as a direct coupling of translator-environment interaction, rather than a process of manipulating structured symbolic mental representations. It shifts focus on non-conscious (priming) and non-representational (dynamic) processes, where the global neuronal workspace may take disambiguating functions of meta-representations or settle for superpositions of heterogeneous wave activation. I will argue that a dynamic principle of relevance is compatible with the Global Neuronal Workspace Theory, and I'll use data from the CRITT TPR-DB to illustrate parameters of such a field of relevance. I will illustrate different trajectories through the relevance field and point out their relation to automatic and reflective translation processes, in support of the Global Neuronal Workspace Theory for Translation.

References:

- Gutt, E. A. (2000). *Translation and relevance: Cognition and context*. Manchester: St. Jerome.
- Dehaene, Stanislas. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*, Penguin
- Krings, Hans. (1986) *Was in den Köpfen von Übersetzern vorgeht*, Narr
- Schaeffer, Moritz and Carl, Michael. 2013. *Shared Representations and the Translation Process: A Recursive Model*. In: *Translation and Interpreting Studies*, Vol. 8, No. 2, 2013, p. 169–190

KEYNOTE

Disciplinary integration and the critical role of sociocognition in theorizing translation

Sandra L. Halverson

For as long as Translation Studies (TS) has claimed some degree of disciplinary autonomy, the field has been concerned with its own relationships with other academic disciplines and the specific forms of intellectual exchange that take place between them. As has often been pointed out, these discussions serve both practical, institutional purposes, as well as epistemological ones (Lambert 2012; Gambier & van Doorslaer 2016). The early claim that TS should be conceived of as an 'interdiscipline' (Snell-Hornby et al. 1994) went hand in hand with the ongoing institutionalization process witnessed throughout the 1980s and 1990s. In contrast, differences soon arose regarding the epistemological objectives of interdisciplinary practices (unification or pluralism/unity or diversity) (Ruano 2006), and this underlying epistemological fracture has been revealed in alternating calls for 'shared ground' (Chesterman & Arrojo 2000) or 'consilience' (Chesterman 2019) as opposed to a pluralistic approach (Seruya et al. 2006; Ruano 2006; Bassnett & Johnston 2019). Further critiques of interdisciplinary practices question the depth of understanding involved (Gile 1999; Shlesinger 2004; Malmkjær 2000), with Gile and Shlesinger referring to 'doorstep interdisciplinarity' in interpreting studies and Ruano referring to an 'uncritical fusion of concepts and methods from different areas of study and paradigms' (2006:49).

Cross-disciplinary relationships continue to be explored and problematized at the level of TS, and the complexity of the concept of interdisciplinarity itself has also been brought into a TS context by Gambier and van Doorslaer (2016). Similar discussions are also taking place within various subdomains, for example CTIS (e.g., Alves 2015), CBTS (de Sutter and Lefer 2019; Kotze 2020, 2022), and sociological TS (e.g., Buzelin & Baraldi 2016). Interdisciplinary relationships remain of vital interest to TS, while at the same time the epistemological status and potential of such relationships are contended.

The TRICKLET symposium invites perspectives on the possibility of theoretical integration within a smaller subdomain of TS, specifically that of linguistically oriented, empirical TS. As such, the call does not address itself to the broad domain of TS, or specifically to questions of interdisciplinarity. It might be suggested, however, that a call for theoretical integration is motivated by the same epistemological impulse as the interdisciplinarity debate, though with a more restricted scope. In grappling with a complex and dynamic object of study such as translation, even theoretical integration in a delimited domain may have implications at a broader level too, depending on the specific theories or concepts being integrated and how the integrative analysis is carried out. It is against this backdrop that this talk may be situated.

In a recent publication, Haidee Kotze and I have argued for a particular foundation for further theorizing in TS, more specifically a usage-based linguistic theory (Halverson & Kotze 2022). In that publication, we demonstrated how this theoretical starting point, and the ontological assumptions on which it builds, allows for a reconceptualizing of important legacy concepts, such as translation norms, and in a principled way renders some other proposed concepts superfluous (in this case Pym's risk management program Pym (2015)). Two important assumptions of this approach are: 1) that translation as it is currently studied in TS is most typically a linguistic or semiotic communicative act, and 2) that language and other conventionalized semiotic systems and communication are fundamentally *sociocognitive* phenomena; that is, phenomena that cannot be reduced to either social or cognitive forces or structures and, consequently, must be theorized in a way that allows for this interwoven and irreducible ontology. In this talk, I extend the argument by exploring how a usage-based theory of language, based on a sociocognitive ontology, can serve as a nexus for interdisciplinary *integration*, in our case, integrating concepts and insights from disciplines focusing on the social or the cognitive, respectively.

In the talk, a brief introduction to a usage-based approach to language is sketched out. Next, an analytical framework for interdisciplinary integration (Repko & Szostak 2021) is introduced, touching on the arguments for and against such integration, and the assumptions on which it rests, with a focus on the specific method presented. This analytical method is then utilized to demonstrate how a usage-based theory of language can incorporate insights from psycholinguistics (the cognitive) and sociology (the social), by virtue of the theory's specific content. Two concepts, one from each field, will be used to articulate a specific integrative process, *structural priming* and *habitus*. After demonstrating how structural priming and habitus can be brought into alignment (integrated) with this theory, the resulting understanding will then be linked back to the concerns of TS, to demonstrate theoretical and methodological consequences. The overarching aim is to show how a usage-based theory of language can serve as a critical nexus for dialogue within TS and between TS and sociology and psycholinguistics, in the interest of theorization and empirical work. The epistemological objective of such an exercise is to facilitate intra- and interdisciplinary dialogue and enrichment, which does not require a universalist aim to be epistemologically useful.

References

- Alves, Fabio. 2015. Translation process research at the interface. Paradigmatic, theoretical, and methodological issues in dialogue with cognitive science, expertise studies and psycholinguistics. In *Psycholinguistic and Cognitive Inquiries into Translation and Interpreting*, Aline Ferreira and John W. Schwieter (eds.), 17-40. Amsterdam: John Benjamins.
- Bassnett, Susan & David Johnston. 2019. The outward turn in Translation Studies. *The Translator* 25 (3): 181-188.

- Buzelin, Hélèn & Claudio Baraldi. Sociology and Translation Studies. In *Border crossings: Translation Studies and other disciplines*, Yves Gambier & Luc van Doorslaer (eds.), 117-139. Amsterdam: John Benjamins.
- Chesterman, Andrew. 2019. Consilience or fragmentation in Translation Studies today? *Slovo.ru: Baltic accent* 10 (1): 9-20.
- Chesterman, Andrew & Rosemary Arrojo. 2000. Shared ground in Translation Studies. *Target* 12 (1): 151-160.
- De Sutter, Gert & Marie-Aude Lefer. 2019. On the need for a new research agenda for corpus-based translation studies: a multi-methodological, multifactorial and interdisciplinary approach. *Perspectives. Studies on translation theory and practice. Perspectives*, 28 (1): 1-23.
- Gambier, Yves & Luc van Doorslaer. 2016. Disciplinary dialogues with translation studies. The background chapter. In *Border crossings: Translation Studies and other disciplines*, Yves Gambier & Luc van Doorslaer (eds.), 1-22. Amsterdam: John Benjamins.
- Gile, Daniel. 1999. Doorstep interdisciplinarity in conference interpreting research. In *Anovar/anosar estudios de traducción e interpretación* (Vol. 1), Alberto Álvarez Lugris & Anxo Fernández Ocampo (eds.), 41-52. Vigo: Universidade de Vigo.
- Halverson, Sandra & Haidee Kotze. 2022. Sociocognitive constructs in translation and interpreting studies (TIS): Do we really need concepts like norms and risk when we have a comprehensive usage-based theory of language? In *Contesting epistemologies in cognitive translation and interpreting studies*, Sandra L. Halverson & Álvaro Marín García (eds.), 51-79. London: Routledge.
- Kotze, Haidee. 2020. Converging what and how to find out why: An outlook on empirical translation studies. In *New empirical perspectives on translation and interpreting*, Lore Vandevoorde, Joke Daems & Bart Defrancq (eds.), 331-370. London: Routledge.
- Kotze, Haidee. 2022. Translation as constrained communication: Principles, concepts and methods. In *Extending the scope of corpus-based translation studies*, Sylviane Granger & Marie-Aude Lefer (eds.), Bloomsbury Academic.
- Lambert, José. 2012. Interdisciplinarity in Translation Studies. In *Handbook of Translation Studies*, Vol. 3, Yves Gambier & Luc van Doorslaer (eds), 81-88. Amsterdam: John Benjamins.
- Malmkjær, Kirsten. 2000. Postscript. Multidisciplinarity in process research. In *Tapping and mapping the processes of translation and interpreting*, Sonja Tirkkonen-Condit & Riitta Jääskeläinen (eds.), 163-170. Amsterdam: John Benjamins.
- Pym, Antony. 2015. Translating as risk management. *Journal of Pragmatics*, 85: 67-80.
- Repko, Allen F. & Rick Szostak. 2021. *Interdisciplinary research. Process and theory*. Los Angeles: Sage.
- Ruano, M. Rosario Martín. 2006. Conciliation of disciplines and paradigms. A challenge and a barrier for future directions in Translation Studies. In *Translation Studies at the interface of disciplines*, Teresa Seruya, Alexandra

- Assis Rosa & João Ferreira Duarte (eds.), 43-53. Amsterdam: John Benjamins.
- Seruya, Teresa, Alexandra Assis Rosa & João Ferreira Duarte. 2006. Introduction. In *Translation Studies at the interface of disciplines*, Teresa Seruya, Alexandra Assis Rosa & João Ferreira Duarte (eds.), 1-6. Amsterdam: John Benjamins.
- Shlesinger, Miriam. 2004. Doorstep interdisciplinarity and beyond. In *Translation research and interpreting research: Traditions, gaps and synergies*, Christina Schaeffner (ed.), 116-123. Amsterdam: John Benjamins.
- Snell-Hornby, Mary, Franz Pöchhacker & Klaus Kaindl. 1994. Preface. In *Translation Studies: An Interdiscipline. Selected papers from the Translation Studies Congress, Vienna, 1992*, Mary Snell-Hornby, Franz Pöchhacker & Klaus Kaindl (eds.), ix. Amsterdam: John Benjamins.

KEYNOTE

Epistemic virtues in (empirical) translation studies

Haidee Kotze
Utrecht University

In this talk I explore how particular intellectual values are infused in the concepts, methods and discourse associated with the subfield of translation research that describes itself as “empirical translation studies”. I do so by drawing on and adapting the notion of “epistemic virtues”, a concept derived from a loosely grouped collection of approaches within epistemology designated as “virtue epistemology” (see Choo 2015; Turri, Alfano & Greco 2021 for overviews). Virtue epistemology concerns itself with questions of how particular intellectual qualities are – and also should be – constructed as virtues and vices in relation to intellectual agents and communities.

My focus is on “empirical translation studies” as an intellectual community, and I attempt to map the way in which this community of practice is collectively constructing intellectual virtues around notions like “empiricism”, “model building”, “theory” and “methodology” – in other words, how these concepts become associated with particular intellectual *values*. This mapping exercise relies on a critical discourse analysis of key texts aligning themselves with “empirical translation studies”, using a corpus-assisted hermeneutic method to identify and explore key words and their associations. Through this exercise, I construct a preliminary set of the epistemic virtues foregrounded in contemporary empirical translation studies.

Subsequent to this, I focus in more detail on one of these epistemic virtues, rigour, and demonstrate how it is constructed in primarily statistical terms. I highlight its relationship with other terms, like “model” and “empirical”, and raise questions (and cautions) about the reification of ostensible virtues (like rigour construed as “sophisticated statistical analysis”) at the expense of attention to other epistemic virtues (see also Freedman 1985, 2010), to the point that some virtues may well tip over into epistemic vices like intellectual and methodological dogmatism and conformity. I conclude by reflecting, in general, on the need for epistemic awareness, care and conscientiousness within (empirical) translation studies.

References

- Choo, Chun Wei. 2015. *The Inquiring Organization: How Organizations Acquire Knowledge and Seek Information*. Oxford Scholarship Online. DOI: 10.1093/acprof:oso/9780199782031.001.0001
- Freedman, David A. 1985. Statistics and the scientific method. In William M. Mason & Stephen E. Fienberg, eds. *Cohort Analysis in Social Research*. New York: Springer. pp. 343-366.
- Freedman, David A. 2010. Statistical models and shoe leather. In Philip B. Stark, Jasjeet Singh Sekhon & David Collier, eds. *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*. Oxford: Oxford University Press. pp. 45-62.

Turri, John, Mark Alfano & John Greco. 2021. Virtue epistemology. In Edward N. Zalta, ed. *The Stanford Encyclopedia of Philosophy*.
<https://plato.stanford.edu/archives/win2021/entries/epistemology-virtue>

KEYNOTE

Top-down and Bottom-Up Approaches to Modeling Translation Variation

Ekaterina Lapshinova-Koltunski
Universität des Saarlandes

We will address methodological issues in modeling translation variation. Generally speaking, there are two approaches to proceed: (1) Top-down: here, we start from linguistic knowledge derived from existing theories and frameworks and define linguistic categories to analyse. With this approach, e.g. variational linguistics (Halliday and Matthiessen, 2014; Biber, 1995 a.o.) or translationese studies (Gellerstam, 1986; Baker, 1995 a.o.), existing theories and frameworks are commonly used as a background to frame research questions and hypotheses, as well as a source for features, i.e. language patterns used in computational modeling. As feature engineering requires much hand-crafted work in this case, such approaches are often applied on a smaller amount of data. Following the second approach, (2) bottom-up, we start from a set of linguistic features of shallow character, e.g. word or part-of-speech n-grams or word embeddings. However, the main requirement for the features is their linguistic interpretability, as the results are interpreted against the background of existing theories and frameworks. Feature engineering is automatic in this case, therefore, a huge amount of data can be employed for such analyses.

Apart from the descriptive representation of translation variation, another aspect that will be addressed in this talk is understanding the mechanisms driving this variation. We want to see to what extent linguistic facets of translation variation can be described and explained in an information-theoretic framework of rational communication (Information Theory, Shannon, 1948).

References

- Mona Baker. 1993. Corpus linguistics and translation studies: Implications and applications. In G. Francis Baker M. and E. Tognini-Bonelli, editors, *Text and Technology: in Honour of John Sinclair*, pages 233-250. Benjamins, Amsterdam.
- Douglas Biber. 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge University Press.
- Martin Gellerstam. 1986. Translationese in Swedish novels translated from English. In L. Wollin and H. Lindquist, editors, *Translation Studies in Scandinavia*, pages 88–95. CWK Gleerup, Lund.
- Claude E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423.

KEYNOTE

The contribution of learner translation corpora to theory development in empirical translation studies: Overview and future prospects

Marie-Aude Lefer
UCLouvain

Corpus-based translation studies (CBTS) has considerably expanded its reach in recent years, with noteworthy advances on many fronts – theoretical, methodological, descriptive – and increased cross-fertilization and interdisciplinarity, both within translation studies and beyond (see De Sutter & Lefer 2020, Gaspari 2022, Kotze 2020 and 2022, Neumann et al. 2022, Vandevoorde et al. 2020). As shown by a recent research survey (Granger & Lefer 2022), CBTS relies strongly on parallel corpora, i.e. corpora containing source texts in a given language, aligned with their translations in another language. These corpora are typically made up of professional or expert translations into the translators' native language (Lefer 2020). Other translation types, such as student translations, non-professional translations, translations into the foreign/second language and machine translation post-editing, remain largely marginal in corpus research. The present talk deals with one of these underexplored corpus resources, namely learner translation corpora (LTC), and aims to show how LTC can contribute to theory development in empirical translation studies. LTC are parallel corpora containing authentic translations produced by translation students or foreign language learners in real-life tasks (e.g. in the translation classroom, as opposed to a laboratory setting; cf. Serbina & Neumann 2021), aligned with their source texts. In addition to representing novice, inexperienced translators, LTC differ from the parallel corpora traditionally used in CBTS in that many of them are *multiple* translation corpora, i.e. they include a number of translations of the same source text, often in large quantities. This unique design offers new opportunities for CBTS and, more generally, empirical translation studies.

The first LTC emerged two decades ago (e.g. PELCRA – Uzar & Walinski 2001, STA – Bowker & Bennison 2003) and were soon followed by similar initiatives (e.g. MISTiC – Castagnoli 2009, MeLLANGE – Castagnoli et al. 2011, NEST – Graedler 2013, UPF – Espunya 2014, RusLTC – Kutuzov & Kunilovskaya 2014, KOPTE – Wurm 2016, CELTraC – Fictumova et al. 2017). In the talk, I will survey recent LTC-based studies in terms of their research objectives, objects of study, key findings and theoretical contributions. I will show that the main foci so far have been on computer-aided translation error analysis and translation quality evaluation (e.g. De Sutter et al. 2017, Espunya 2014, Kübler et al. 2022, Kunilovskaya & Lapshinova-Koltunski 2019, Vela et al. 2014), mostly with a view to informing translator training. Alongside this core applied-research strand, new types of LTC-based empirical investigation, with a somewhat stronger focus on theorization, have started to emerge in recent years. This is true, for instance, of the study of translation features such as explicitation, normalization and levelling-out, which had hitherto been approached mainly through the lens of expert translation in CBTS

(see e.g. Castagnoli 2016, Kunilovskaya et al. 2018, Lapshinova-Koltunski 2022, Looock 2020, Redelinguys 2016, Redelinguys & Kruger 2015). I will also discuss LTC-based studies that have addressed key constructs such as translation competence acquisition (e.g. through longitudinal corpus studies; Wurm 2020) and translation variation (Castagnoli 2020).

The next part of my talk will be devoted to theory-oriented research based on the *Multilingual Student Translation* (MUST) corpus. MUST is an international LTC collection initiative that brings together 45 partner teams worldwide (Granger & Lefer 2020). The corpus currently comprises ca 400 source texts (ranging from 150 to 1,000 words in length) and 6,000 student translations produced by ca 3,000 students, with 19 languages represented. It includes numerous text types, both general (news and opinion articles, excerpts from novels, etc.) and specialized (financial reports, tourist guides, instruction manuals, contracts, etc.). The strengths of the MUST corpus include its rich standardized metadata relating to the source texts, translation tasks and learners (40+ metadata rubrics), and the Translation-oriented Annotation System (TAS) developed collaboratively within the MUST network to support research on translation quality across language pairs (Granger & Lefer 2021). In the talk, I will zoom in on two topics currently being investigated at UCLouvain in connection with the English-French pair, namely (1) simplification phenomena in student translations and their sources of explanation, tapping into the rich learner-related metadata collected in the MUST project and insights from contrastive linguistics (Penha-Marion et al. forthcoming) and (2) individual variation and translation invariance, relying on a combination of corpus-based and corpus-driven linguistically informed analyses (pre-selected source-text features, such as figurative language, cognates, established/ad hoc noun sequences; target-text n-grams and POS-grams).

The talk will end with a look to the future. I will take stock of the LTC-derived theoretical advances made so far to sketch key avenues for LTC-based research, such as the link between translation variation, quality and literality, while also paying attention to the most acute challenges that lie ahead.

References

- Bowker, L. & Bennison, P. (2003). Student Translation Archive: Design, development and application. In F. Zanettin, S. Bernardini & D. Stewart (eds.) *Corpora in Translator Education*. London & New York: Routledge, 103-117.
- Castagnoli, S. (2020). Translation choices compared: Investigating variation in a learner translation corpus. In S. Granger & M.-A. Lefer (eds.) *Translating and Comparing Languages: Corpus-based Insights*. Corpora and Language in Use Proceedings 6. Louvain-la-Neuve: Presses universitaires de Louvain, 25-44.
- Castagnoli, S. (2016). Investigating trainee translators' contrastive pragmalinguistic competence: A corpus-based analysis of interclausal linkage in learner translations. *The Interpreter and Translator Trainer* 10(3), 343-363.
- Castagnoli, S. (2009). *Regularities and variations in learner translations: A corpus-based study of conjunctive explicitation*. PhD thesis, Pisa University.
- Castagnoli, S., Ciobanu, D., Kübler, N., Kunz, K. & Volanschi, A. (2011). Designing a Learner Translator Corpus for Training Purposes. In N. Kübler (ed.) *Corpora*,

- Language, Teaching, and Resources: From Theory to Practice*. Bern: Peter Lang, 221-248.
- De Sutter, G., Cappelle, B., De Clercq, O., Loock, R., & Plevoets, K. (2017). Towards a corpus-based, statistical approach to translation quality: Measuring and visualizing linguistic deviance in student translation. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 16, 25-39.
- De Sutter, G. & Lefer, M.-A. (2020). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives* 28(1), 1-23.
- Espunya, A. (2014). The UPF learner translation corpus as a resource for translator training. *Language Resources and Evaluation* 48, 33-43.
- Fictumova, J., Obrusnik, A., & Stepankova, K. (2017). Teaching specialized translation error-tagged translation learner corpora. *Sendebear* 28, 209-241.
- Gaspari, F. (2022). Expanding the reach of corpus-based translation studies: The opportunities that lie ahead. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. London: Bloomsbury, 42-63.
- Graedler, A.-L. (2013). NEST—A corpus in the brooding box. *Studies in Variation, Contacts and Change in English*, 13.
- Granger, S. & Lefer, M.-A. (2022). Corpus-based translation and interpreting studies: A forward-looking review. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. London: Bloomsbury, 13-41.
- Granger, S. & Lefer, M.-A. (2021). *Translation-oriented Annotation System manual (Version 2.0)*. CECL Papers 3. Louvain-la-Neuve: Centre for English Corpus Linguistics/Université catholique de Louvain. Available online: <https://uclouvain.be/en/research-institutes/ilc/cecl/cecl-papers.html>
- Granger, S. & Lefer, M.-A. (2020). The *Multilingual Student Translation* corpus: A resource for translation teaching and research. *Language Resources and Evaluation* 54, 1183-1199.
- Kotze, H. (2022). Translation as constrained communication: Principles, concepts and methods. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. London: Bloomsbury, 67-97.
- Kotze, H. (2020). Converging what and how to find out why: An outlook on empirical translation studies. In L. Vandevoorde, J. Daems & B. Defrancq (eds.) *New Empirical Perspectives on Translation and Interpreting*. Abingdon: Routledge, 333-370.
- Kübler, N., Mestivier, A., & Pecman, M. (2022). Using comparable corpora for translating and post-editing complex noun phrases in specialized texts: Insights from English-to-French specialized translation. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. London: Bloomsbury, 237-266.
- Kunilovskaya, M. & Lapshinova-Koltunski, E. (2019). Translationese features as indicators of quality in English-Russian human translation. In I. Temnikova, C. Orasan, G. Corpas Pastor & R. Mitkov (eds.) *Proceedings of the 2nd Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT 2019)*, 47-56.

- Kunilovskaya, M., Morgoun, N., & Pariy, A. (2018). Learner vs. professional translations into Russian: Lexical profiles. *Translation and Interpreting* 10(1), 33-52.
- Kutuzov, A., & Kunilovskaya, M. (2014). Russian learner translator corpus: Design, research potential and applications. In P. Sojka, A. Horak, I. Kopecek & K. Palak (eds.) *Text, Speech and Dialogue. Lecture Notes in Computer Science*. Berlin: Springer, 315-323.
- Lapshinova-Koltunski, E. (2022). Detecting normalisation and shining-through in novice and professional translations. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. Bloomsbury, 182-206.
- Lefer, M.-A. (2020). Parallel corpora. In M. Paquot & S. Th. Gries (eds.) *A Practical Handbook of Corpus Linguistics*. Cham: Springer, 257-282.
- Loock, R. (2020). It's non-canonical word order that you should use! A corpus approach to avoiding standardized word order in translated French. In S. Granger & M.-A. Lefer (eds.) *Translating and Comparing Languages: Corpus-based Insights*. Corpora and Language in Use Proceedings 6. Louvain-la-Neuve: Presses universitaires de Louvain, 69-85.
- Neumann, S., Freiwald, J., & Heilmann, A. (2022). On the use of multiple methods in empirical translation studies: A combined corpus and experimental analysis of subject identifiability in English and German. In S. Granger & M.-A. Lefer (eds.) *Extending the Scope of Corpus-based Translation Studies*. London: Bloomsbury, 98-129.
- Penha-Marion, L. A. de S., Gilquin, G., & Lefer, M.-A. (forthcoming). The effect of directionality on lexico-syntactic simplification in French><English student translation. In H. Kotze & B. Van Rooy (eds.) *Constraints on Language Variation and Change in Complex Multilingual Contact Settings*. Amsterdam: John Benjamins.
- Redelinghuys, K. (2016). Levelling-out and register variation in the translations of experienced and inexperienced translators: A corpus-based study. *Stellenbosch Papers in Linguistics* 45, 189-220.
- Redelinghuys, K. & Kruger, H. (2015). Using the features of translated language to investigate translation expertise: A corpus-based study. *International Journal of Corpus Linguistics* 20(3), 293-325.
- Serbina, T. & Neumann, S. (2021). Translation product and process data: A happy marriage or worlds apart? In S. L. Halverson & I. Marín García (eds.) *Contesting Epistemologies in Cognitive Translation and Interpreting Studies*. Abingdon: Routledge, 131-152.
- Uzar, R. & Walinski, J. (2001). Analysing the fluency of translators. *International Journal of Corpus Linguistics* 6, 155-166.
- Vandevoorde, L., Daems, J., & Defrancq, B. (eds.) (2020). *New Empirical Perspectives on Translation and Interpreting*. Abingdon: Routledge.
- Vela, M., Schumann, A.-K., & Wurm, A. (2014). Beyond linguistic equivalence. An empirical study of translation evaluation in a translation learner corpus. In *Proceedings of the EAFL 2014 Workshop on Humans and Computer-assisted Translation*, 47-56.

- Wurm, A. (2020). Translation quality in an error-annotated translation learner corpus. In S. Granger & M.-A. Lefer (eds.) *Translating and Comparing Languages: Corpus-based Insights*. Corpora and Language in Use Proceedings 6. Louvain-la-Neuve: Presses universitaires de Louvain, 141-162.
- Wurm, A. (2016). Presentation of the KOPTÉ Corpus and Research Project. https://www.academia.edu/24012369/Presentation_of_the_KOPTÉ_Corpus_and_Research_Project.

KEYNOTE

Aeons later

Moritz Schaeffer
Johannes Gutenberg-Universität Mainz

Studies investigating the cognitive processes during translation aim at modelling the temporal dimensions of the translation process in terms of the architectural aspects which lead from input to output – disregarding to a very large extent how this output might be received. How translations are or might be received has received considerable attention in the history of the discipline. The implied target reader has featured prominently in many theoretical accounts of translation but has not been studied in conjunction with the processes that occur on a millisecond scale inside the translator's skull. Chesterman (2000) calls for a causal model of translation, because only a "...causal model explicitly makes it possible to posit explanatory predictions." (25) This talk will ask whether and if so, how it is possible to establish a causal dependency between what happens in the first 200ms of contact with a source item and its eventual reception in a target language aeons later. This talk will attempt to answer why such an explanatory prediction might be useful by way of attempting to chart the requirements for such an answer.

References

Chesterman, A., 2000. A Causal Model for Translation Studies, in: Olohan, M. (Ed.), *Intercultural Faultlines. Research Models in Translation Studies I. Textual and Cognitive Aspects*. St. Jerome, Manchester, pp. 15–27.

KEYNOTE

Is there still a role to play for corpora when translation theories turn cognitive?

Gert De Sutter
Ghent University

It is widely acknowledged that observational and interventionist research in linguistically-oriented descriptive translation studies has produced an increasingly fine-grained insight into the way language is used in translation, and how this is linked to user-, text- and context-related variables (see, e.g., Vandevoorde, Daems & Defrancq 2020, Granger & Lefer 2022 for a recent overview). This could only be achieved by significant methodological and analytical sophistication, which included key-stroke logging, eye tracking, context-sensitive corpora, confirmatory and exploratory statistics, etc. These empirical advances have recently been accompanied by several attempts to integrate these findings in an encompassing and coherent theory which takes into account insights from related disciplines such as psycholinguistics, second language acquisition and register studies (to name but a few). One important source of information for such a theoretical endeavor is cognitive in nature, inspired by models in psycholinguistics and cognitive linguistics, among others. Notable illustrations of this include the recursive model of translation developed by Carl & Schaeffer (2013) and Halverson's Gravitational Pull Hypothesis (2017).

Against this background of cognitive orientation, one might wonder what role corpus data still have to play - if any at all - in shedding light on the cognitive underpinnings of the translation act. In other words, can corpora, which amalgamate utterances representing only the final stage of a complex communication process, produced by a wide variety of translators who have different levels of experience and expertise, different language and educational backgrounds, and work under different circumstances, still reveal something about cognitive processing during translation? In this talk I will argue that *some* patterns found in *qualitative* corpus data, when analyzed *with care*, can still be seen as reflections of cognitive processing and can thus inform research designs using more direct, interventionist methods, ultimately leading to a stable, coherent theory of language use in translation. I will illustrate this by means of a case study on subject placement in Dutch which is based on a controlled corpus of expert and student translations, combining insights from psycholinguistics, cognitive linguistics and language typology.

References

- Granger, S. & M.A. Lefer (eds.) (2022). Extending the scope of corpus-based translation studies. Bloomsbury.
- Halverson, S.L. (2017). Gravitational pull in translation. Testing a revised model. In G. De Sutter, M.-A. Lefer & I. Delaere (eds), Empirical Translation Studies.

New Methodological and Theoretical Traditions. Berlin: Mouton de Gruyter, 9-46.

Schaeffer, M. & M. Carl (2013). Shared representations and the translation process: A recursive model. *Translation and Interpreting Studies* 8 (2), 169-190.

Vandevoorde, L., J. Daems & B. Defrancq (eds.) (2020). *New empirical perspectives on translation and interpreting*. Routledge.

Paper Abstracts

Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters

Nannan Liu
The University of Hong Kong

The professional interpreter speaks in the first-person singular (Harris, 1990). Previous research has reported shifts from this norm towards first-person plural to be the most frequent pronoun shift in political institutional interpreting, such as in the European Parliament (EP; Beaton-Thome, 2010) and Chinese premier press conferences (CPPCs; Gu & Tipton, 2020). First-person pronoun (FPP) shifts have received social, cognitive, and linguistic explanations. First, institutional interpreters' FPP "plural preference" was viewed as having been prompted by (i) topics that encode "ideological salience," e.g. the Israel–Palestine relationship (Mona-celli, 2009, pp. 97–98) or Guantánamo Bay detainees (Beaton-Thome, 2010, 2013), and (ii) institutional loyalty (Beaton-Thome, 2010; Fu & Chen, 2019). Prosodic stress and disfluencies were argued to reveal interpreter ideological positioning (Beaton-Thome, 2013). Second, cognitive demands affect an interpreter's preferences for the surface forms and underlying referents of FPPs. In terms of forms, Plevoets and Defrancq (2018, p. 9) identified *il nous faut* ("we need to") as a formulaic sequence in EP French source speech. Li and Halverson (2020) demonstrated the association between self- and other-priming and CPPC interpreters' selection of FPP sequences. In terms of underlying referents, shifts towards plurals indexing dominant institutions can be explained with reference to the bondedness hierarchy (Croft, 2002) and performance–grammar correspondence hypothesis (Hawkins, 2004). Third, languages such as Mandarin and Spanish allow zero subjects and so when interpreting into subject-obligatory languages, some shifts towards FPPs are necessary.

This paper pursues two goals: (i) identifying the factors that influence political institutional interpreters' choices between first-person singular and plural, and (ii) disentangling the main explanations proposed for institutional interpreters' first-person "plural preference." I extend Behavioural Profile and Multifactorial Prediction and Deviation Analysis using Regression/Random Forests (MuPDAR[F]; Gries, 2010; Gries & Deshors, 2014) to achieve the goals.

This study adopts the usage-based approach, under the assumption that first-person choices derive from domain-general cognitive processes such as chunking (Bybee, 2010). Chunking effects are evidenced by a rise in the frequency of co-occurrence, an association with repetition/priming, production fluency, morphosyntactic bondedness, and phonetic erosion (Bybee, 2010, pp. 34–56). Chunking propels grammaticalisation "in specific contexts and constructions" (Narrog&Heine, 2021, p. 1). Because of the underlying role of chunking in grammaticalisation, the same diagnostics for chunking also apply to grammaticalisation (Bybee, 2010; Narrog & Heine, 2021).

In light of the literature and theoretical models, this study examines the following hypotheses.

1. By the institutional alignment account,

- a) the first-person plural would increase in frequency, attract stress or disfluencies, and be less grammaticalised vis-à-vis the singular when speakers/interpreters signal alignment with institutional positions on ideologically salient topics across original Chinese (OC), interpreted English (IE), and native English (NE);
- b) the plural in IE would be less grammaticalised than that in OC as interpreters had strengthened institutional alignment on ideologically salient topics compared with source speakers (Beaton-Thome, 2010; Fu & Chen, 2019; Gu & Tipton, 2020);

2. By the cognitive processing explanations,

- a) plural constructions would be more grammaticalised than the singular across the three varieties when speakers process and produce complex forms and referents and interpreters do so in the source and target;
- b) plural constructions in IE would be more grammaticalised than those in OC and NE owing to the greater cognitive demands of interpreting;

3. By the linguistic specificity account, plural constructions would be more grammaticalised than the singular in IE when there is no subject form in the source segment, because interpreters may have developed a plural formula for dealing with zero-subject inputs. Hypothesis 3 aligns with Hypothesis 2.

FPP cases were extracted from parallel CPPC interpreting and comparable speech corpora (Author, 2020). FPP cases in OC and consecutive IE were sampled, with IE data made by five different staff interpreters of the Chinese Ministry of Foreign Affairs, who worked into their B language. NE cases were obtained from the State of the Union addresses of former US president George W. Bush. The motivation behind selecting Bush's addresses is to minimise the differences between the parallel and comparable corpora in mode (read-out from the teleprompter, scripts, or consecutive notes), timeframe (in the 2000s–2010s), duration (50–60 minutes of native-language speech delivery per session), register (Liu, 2021), functions (policy debriefing, support rallying, and image management; Chen, 2007; Kreiser & Greene, 2020), and size. Multifactorial studies assume that all cases are freely variable (Kruger & De Sutter, 2018, p. 256) between the singular and plural, and cases that impose an FPP form were discarded.

The dataset comprises 2,438 cases and were tagged for 33 variables. The variable topic is related to Hypothesis 1; the topics of Taiwan, the Iraq War, and international relations are considered “ideologically salient” and may favour the plural. Cognitive and linguistic variables that are related to Hypotheses 2 and 3 include self- and other-priming, and those assumed to reflect the complexity of the grammatical context (e.g. aspect, clause type) and underlying referents (e.g. bondedness, group). Other-priming relates both to cognitive processing (in cases of the FPP) and linguistic specificities (in cases of zero). Prosodic variables are related to Hypotheses 1–3 and differentiate the degree of phonetic erosion in FPP constructions.

The MuPDAR(F) approach was chosen over interaction models because the latter face more serious complete separation and multicollinearity issues, which means that variables such as topic that are essential to disentangle the explanations had to be discarded in the process of finding the optimal interaction models. MuPDAR(F) was first used to test Hypothesis 1 and part of Hypothesis 2 about complex forms in production. Because of complete separation issues, MuPDAR(F) cannot address how complex forms in the source and complex referents (Hypothesis 2), and zero-subject inputs (Hypothesis 3) bear on interpreter “plural preference.” Therefore, two more regressions were fitted.

The results support Hypotheses 1a and 2a in OC and NE, 2 and 3 in IE, reject 1b, and largely discredit 1a in IE. The analyses show extensive grammaticalisation of plural constructions in the interpreted speech, relative to those in source and comparable speech with the singular as the reference, relative to those in non-interpreting, and relative to interpreters’ singulars. Institutional interpreters are found to make the “plural preference” in compliance with conventions of political speech, and when processing and producing complex forms and referents in the source and target, and zero-subject inputs. The institutional alignment explanation proposed in previous studies fails to account for the extensive grammaticalisation of plural constructions in interpreted speech.

This study provides robust methods for context analyses and cross-linguistic prosodic transcription, and exemplifies the relevance of usage-based, multifactorial designs for interpreting research.

References

- Author. (2020). *Title of defended thesis* (Doctoral dissertation).
- Beaton-Thome, M. (2010). Negotiating Identities in the European Parliament: The Role of Simultaneous Interpreting. In M. Baker, M. Olohan, & M. Calzada-Pérez (Eds.), *Text and Context: Essays on Translation and Interpreting in Honour of Ian Mason* (pp. 117–138). Manchester: St. Jerome.
- Beaton-Thome, M. (2013). Whats in a word? Your enemy combatant is my refugee: The role of simultaneous interpreters in negotiating the lexis of Guantánamo in the European Parliament. *Journal of Language and Politics*, 12(3), 378–399.
- Bybee, J. L. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Chen, L. (2007). *Wenhua yujing yu zhengzhi huayu: Zhengfu xinwen fabuhui de huayu yan-jiu [Cultural context and political discourse: A discourse analysis of government press conferences]* (PhD Thesis, Shanghai International Studies University, Shanghai, China).
- Croft, W. (2002). *Typology and Universals* (2nd ed.). Cambridge: Cambridge University Press.
- Fu, R., & Chen, J. (2019). Negotiating interpersonal relations in Chinese-English diplomatic interpreting. *Interpreting*, 21(1), 12–35.
- Gries, S. T. (2010). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon*, 5(3), 323–346.
- Gries, S. T., & Deshors, S. (2014). Using regressions to explore deviations between interlanguage and native language: Two suggestions. *Corpora*, 9(1), 109–136.

- Gu, C., & Tipton, R. (2020). (Re-)voicing Beijings discourse through self-referentiality: A corpus-based CDA analysis of government interpreters discursive mediation at Chinas political press conferences (19982017). *Perspectives*, 28(3), 406–423.
- Harris, B. (1990). Norms in Interpretation. *Target*, 2(1), 115–119.
- Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.
- Kreiser, M., & Greene, M. (2020). *History, Evolution, and Practices of the Presidents State of the Union Address: Frequently Asked Questions* (tech. rep. No. R44770). Congressional Research Service. Washington, DC.
- Kruger, H., & De Sutter, G. (2018). Alternations in contact and non-contact varieties: Reconceptualising that-omission in translated and non-translated English using the MuPDAR approach. *Translation, Cognition & Behavior*, 1(2), 251–290.
- Li, Y., & Halverson, S. L. (2020). A corpus-based exploration into lexical bundles in interpreting. *Across Languages and Cultures*, 21(1), 1–22.
- Liu, N. (2021). Register shifts in political conference interpreting: A multidimensional analysis. In J. Pan, S. L. Halverson, & J. Munday (Eds.), *Translating and Interpreting Political Discourse: New Trends and Perspectives*. Leiden: Brill.
- Monacelli, C. (2009). *Self-Preservation in Simultaneous Interpreting*. Amsterdam: John Benjamins.
- Narrog, H., & Heine, B. (2021). *Grammaticalization*. Oxford: Oxford University Press.
- Plevoets, K., & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency uh(m). *Interpreting*, 20(1), 1–28.

Embodied Metarepresentations

Nicolás Araneda Hinrichs^{1, 2}, Maryam Foradi¹, Tariq Yousef¹, Susanne Triesch^{3, 4} & Elisa Hartmann⁵

¹University of Leipzig, ²University of Concepción, ³Heinrich Heine University Düsseldorf, ⁴University of Cologne, ⁵Chemnitz University of Technology

Meaning has been established pervasively as a central concept throughout the disciplines that were involved in the cognitive revolution (Pylyshyn, 1984). Interlinguistic phenomena constitute one of the most overt displays of the shared representational relationships (Semin & Smith, 2007) that give rise to it. Despite the instability of meaning in relation to its underdetermination by reference (Baumgarten, 2012), it can be a *tertium comparationis* or “third comparator” for extended human cognition (Hatim & Munday, 2019), if gauged through invariants that exist in transfer processes such as translation, as representational content emerges in interaction (Meteyard et al., 2012) and, moreover, all languages and cultures are rooted in pan-human experience, and thus share and express species-specific ontology.

Furthermore, seen as a cognitive competence (Barsalou, 1992), meaning doesn't stop outside of the body but dynamically depends and couples with other agents and the environment (Enfield & Kockelman, 2017; Brette, 2018; Rojas-Líbano & Parada, 2020), and can be considered to be embedded, extended, embodied and enacted (Muñoz Martín, 2017). A novel approach to exploring the transfer properties of semantic frames (SF) is presented; these are organized according to the principles of frame semantics (FS) in the Berkeley online database FrameNet (FN) (Fillmore et al., 2003; Petruck, 1997; Ruppenhofer et al., 2016). In translation and interpreting studies (TIS), SF are taken to be at the core of the meaning that is carried across. Čulo (2017) has proposed a primacy of frame model building on “the idea that preserving the conceptual information connected with a frame in the source language by picking an adequate frame in the target language is a core procedure in translation” (Čulo 2013: 144). While picking an adequate frame will in many cases mean choosing the maximally comparable frame to the one in the source text, a number of factors can override this principle and bring about frame shifts (Čulo 2017: 479). In order to explore the characteristics of these, those SF evoked by the natural semantic metalanguage's (NSM) primes/primitives SEE and FEEL (Wierzbicka, 1996; Goddard, 2010)- were extracted from the European Parliament Proceedings Parallel Corpus (EuroParl), a parallel corpus (Leech, 1991) that allows for the automatic word alignment of the items with their translations (Koehn, 2005). EuroParl was chosen in light of its “(...) free availability, size, linguistic diversity, data authenticity (...)” (Ustaszewski, 2019: 107), which makes it ideal for translation-oriented corpus-based inquiries, moreover if applied as a data-driven approach that serves to characterize mental or sociocultural aspects of interlingual phenomena.

To achieve the above-mentioned aim of study the method of semantic mirroring (SM) (Vandevoorde, 2020) was chosen and thus several fully- and semi-automated steps were taken. Firstly, given that EuroParl is a database which consists of sentences that were either uttered in an L1 or translated from any of the 24 official

languages of the European Union, the aligned sentences' meta-data was used to filter out the ones for which it wasn't specified whether their L1 was English. After extracting all the sentences with both English as its source language and which also contained conjugated forms of the above-named primitives -that is, SEE and FEEL-, the original sentences were aligned with their German and Spanish translations at word level using Multilingual BERT (Nagata et al., 2020). In this way, it was observed how each token was translated into German and Spanish. Due to the authentic nature of the translations performed by interpreters of the European Parliament, it is evident that not all the instances of both verbs have a one-to-one correspondence; thus, the data had to be further sorted by filtering out those sentences, for which there was no clear corresponding in the target language; for instance, a frequent case consisted of a verb being aligned with a NULL value or a preposition.

Having approximately 11,000 sentences for the language pair EN-DE and 9,000 sentences for EN-ES, DeepL was used for their back-translation into English so as to achieve SM as described by Vandevoorde (2020). The compared examination of SF evoked by each verb was facilitated by the use of Large Ontology Multilingual Extraction (LOME) (Xia et al., 2021), a tool for multilingual frame-semantic annotation, in order to tag each verb in original English, its translation in German and Spanish, its back-translations into English with the SF activated by the verb in each sentence per language. This concluded the semi-automatic query in search of prototypical patterns of translation text production (Alves and Vale, 2009) and further statistical analysis were performed. The present paper offers results and, additionally, experimental designs which are proposed for the neurophysiological examination of the observed universals' (Baker, 2018) correlates (García, 2019), on the basis of this initial exploratory approach, in order to combine both methods into an empirically-based model of translation.

References

- Alves, F. & Vale, D. (2017). On drafting and revision in translation: A corpus linguistics oriented analysis of translation process data. In Hansen-Schirra, S., Neumann, S. & Čulo, O. (Eds.), *Annotation, Exploitation and Evaluation of Parallel Corpora*. Language Science Press.
- Baker, M. (2018). *In other words: A coursebook on translation*. Routledge.
- Barsalou, L.W. 1992. Frames, Concepts, and Conceptual Fields. In A. Lehrer & E. F. Kittay (eds.), *Frames, Velds, and contrasts*, 21–74. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behav. Brain. Sci.* 42:49
- Čulo, O. (2013). Constructions-and-frames analysis of translations: the interplay of syntax and semantics in translations between english and german. *Constructions and Frames*, 5(2), 143–167. <https://doi.org/10.1075/cf.5.2.02cul>
- Czulo, O. (2017). Aspects of a primacy of frame model of translation. In S. Hansen-Schirra & O. Czulo (Eds.), *Translation and multilingual natural language processing: Vol. 7. Empirical modelling of translation and interpreting* (pp. 465–490). Freie Univ. Berlin, FB Philosophie und Geisteswissenschaften.

- Enfield, N. J. & Kockelman, P. (eds.) 2017. *Distributed Agency*. Oxford: Oxford University Press. FrameNet (Ed.). Framenet data: frame index. International Computer Science Institute, Berkeley. <https://framenet.icsi.berkeley.edu/fndrupal/frameIndex>
- Fillmore, C. J., Johnson, C. R. & Petruck, M. R.L. (2003). Background to FrameNet. *International Journal of Lexicography* 16.3: 235–250. Goddard, C. (2010). The Natural Semantic Metalanguage approach. In Bernd Heine and Heiko Narrog (eds.) *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press. 459-484.
- Hatim, B. & Munday, J. (2019). *Translation: An advanced resource book*. Routledge.
- Koehn, P. (2005). Europarl: A parallel corpus for statistical machine translation. In *Proceedings of the 10th machine translation summit* (pp. 79-86). Asia-Pacific Association for Machine Translation.
- Leech, G. (1991). The state of the art in corpus linguistics. In Aijmer, K. & Altenberg, B. (Eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 8-29. Longman.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788-804.
- Muñoz Martín, R. (2017). Looking toward the future of cognitive translation studies. In Schwieter, J. & Ferreira, A. (Eds.), *The Handbook of Translation and Cognition* (pp. 555-572). Wiley & Sons.
- Petruck, M. R. L. (1997). Frame semantics. In J. Blommaert, C. Bulcaen, J.-O. Östman, J. Verschueren, & E. Versluys (Eds.), *Handbook of pragmatics* (pp. 1–8). Benjamins.
- Pylshyn, Z. (1984). *Computation and Cognition*. Cambridge: MIT Press.
- Rojas-Líbano, D. & Parada, F. J. (2020). Body-world coupling, sensorimotor mechanisms and the ontogeny of social cognition. *Frontiers in Psychology* 10: 3005. doi: 10.3389/fpsyg.2019.03005
- Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., Baker, C. F., & Scheffczyk, J. (2016). *FrameNet II: Extended Theory and Practice*. (Revised November 1, 2016).
- Semin, G. R. & Smith, E. R. (2007). Situated Social Cognition. *Current Directions in Psychological Science* (16: 3): 132-135.
- Ustaszewski, M. (2019). Optimising the Europarl corpus for translation studies with the EuroparlExtract toolkit. *Perspectives*, 27:1, 107-123.
- Vandevoorde, L. (2020). Semantic differences in translation : exploring the field of inchoativity. *Language science press*. Wierzbicka, A. (1996). *Semantics: Primes and universals: Primes and universals*. Oxford University Press.
- Xia, P., Qin, G., Vashishtha, S., Chen, Y., Chen, T., May, C., Harman, C., Rawlins, K., White, A. S., & van Durme, B. (2021, March 15). LOME: Large Ontology Multilingual Extraction. <http://arxiv.org/pdf/2101.12175v2>

Building a communication-based analytical framework in situated translation training

Jun Yang
Swansea University

The paper reports an empirical study on analysing students' collaboration in a large multilingual localisation project with 77 students in three MA programmes working together in 12 language pairs on content in a variety of file formats. It discusses the communication across cultural barriers and time zones, the planning and organisation of localisation projects, the influence of different roles for the collaborative localisation projects, and the dynamic relationship between freelancer, LSP (language service provider) and end-client.

While the titles of roles in translation and localisation projects are well-known - project manager, translator, reviser, terminologist, DTP expert - their level of involvement in different stages of a project is not. For instance, not many freelance translators are aware of the extended effort and time put in by project managers (PMs) who lead complex multilingual projects. This study shows that PMs' workload can outweigh that of individual linguists tenfold. Moreover, little is known about which specific communicative interactions between different actors in collaborative translation projects are instrumental to their success. However, when the quality of the final product depends on both the quality of the individual work, and the quality and efficiency of the interactions between the project participants, remaining ignorant of the project big picture, and particularly of the impact of one's performance on the remaining project stages and the rest of the team, often leads to project failure.

Situated learning which refers to 'a context-dependent approach under which learners are exposed to real-life and/or highly simulated work environments and tasks, both inside and outside the classrooms' is often used in translation training programmes (González-Davies & Enríquez-Raído, 2016:1). Simulations of translation/localisation projects enable students to experience different professional roles and to use a range of tools and resources to work on authentic source texts, in which the ultimate goal is to help students construct a comprehensive picture of how translation/localisation is done in the professional world. Assessment of such training approach often focuses on students' technological competence and/or linguistic and translation competence, while students' performance in other aspects, such as service provision, personal and interpersonal communication, is often overlooked. This study will demonstrate how to use the communicative interactions among the students as a supplement measurement for their performance throughout the collaborative localisation project. The interaction discourse and communication strategy can reflect students' acquisition of non-linguistic competences as defined in the European Master's in Translation Network (2017) competence framework.

In this study, students were encouraged to communicate with each other regarding the localisation project as a whole, as well as specific translation challenges which they encountered during the project – 2,295 email exchanges and 126 forum posts were collected and analysed.

Informed by the research on the role of communication in facilitating collaborative learning (Soller, 2001), and the use of collaborative environments in community building and in scaffolding translator trainee's experience (Babych et al., 2012), an interaction typology that caters for both professional and learning activities in collaborative localisation projects has been developed. The interaction typology was used for Dialogue Act Analysis (DAA) (Jordan & Henderson, 1995) to find out: the communication patterns in students' collaborative localisation project; the frequency and category of interaction in different stages of the project; and the peer-learning process embedded in communication. Social Network Analysis (SNA) (Borgatti, et al., 2009) was then used to visualise the communication patterns along with the project's key stages, which enhances the understanding of professional roles and workflows in the collaborative environment.

Using the data collected from online asynchronous collaboration, this empirical study identified the characteristics of roles and communicative patterns that form the collaboration among multiple parties and highlighted individuals' roles and influence in a dynamic collaborative network. It will showcase how a communication-based analytical framework is built and used, which will offer insights in conducting empirical research with cross-disciplinary methods. Results of the study will also benefit translator training programmes by enabling them to devise realistic tasks based on an empirically-proven picture of collaborative workflows and practices.

References

- Babych, B., Hartley, A., Kageura, K., Thomas, M., Utiyama, M., 2012, MNH-TT: a collaborative platform for translator training. *Translating and the Computer*, 34, 29-30, November 2012, London.
- Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G., 2009, Network Analysis in the Social Sciences. *Science*, 323, 892–895.
- European Master's in Translation Board, 2017, Competence Framework 2017. Retrieved from https://ec.europa.eu/info/sites/default/files/emt_competence_fwk_2017_en_web.pdf
- González-Davies, M., Enríquez-Raído, V., 2016, Situated learning in translator and interpreter training: bridging research and good practice, *The Interpreter and Translator Trainer*, 10:1, 1-11
- Jordan, B., Henderson, A., 1995. Interaction Analysis: Foundations and Practice. *Journal of the Learning Sciences*, 4, 39–103.
- Soller, A., 2001. Supporting social interaction in an intelligent collaborative learning system. *International Journal of Artificial Intelligence in Education*, 12, 40–62.

Syntactic characterization of constrained language based on dependency treebanks of translated English and non-native English

Ruimin Ma, Yue Jiang
Xi'an Jiaotong University

As a product of a code-switching process with bilingual activation, translated language is assumed to have universal features inherent in the translation process (Baker, 1993: 245) that are distinctive from its monolingual counterpart. However, are these alleged universal features translation-specific/intrinsic or are they shared by other types of bilingualism-influenced constrained language use such as non-native/learner language, edited language and other contact language varieties? The pondering over this question has gradually shifted the discussion of translation universals (TU) (Baker, 1993) towards a more broad and unified theoretical paradigm: constrained communication and language mediation (Lanstyák & Heltai, 2012; Kruger & van Rooy, 2016; Kruger, 2018; Granger & Lefer, 2020; Kotze, 2020). So far, the empirical quest for shared features of mediated/contact/constrained language varieties, still thin on the ground (Gaspari & Bernardini, 2010; Kruger & van Rooy, 2016a; Rabinovich et al., 2016; Ferraresi, 2019; De Sutter & Lefer, 2020; Ivaska & Bernardini, 2020; Kajzer-Wietrzny & Ivaska, 2020; Kajzer-Wietrzny & Grabowski, 2021), has dealt mainly with Indo-European language pairs and single-faceted local lexical and syntactic features. Syntactic indices such as collocations, complementizer *that*, mean clause/sentence length, n-grams, cohesive markers, etc., with their length or frequency have been examined statically and isolatedly. On the other hand, as similar cognitive constraints involved in bilingual activation are at the core of what we presuppose and account for the commonalities between translation and non-native production, more sophisticated syntactic metrics with cognitive dimension remain to be explored to better capture the underlying cognitive mechanisms and the nature of constrained language varieties. Regarding the methodological issues in research design, Kruger's (2018) constraint dimension framework (also see Kotze, 2020), has scarcely or insufficiently been taken into account in previous studies, thus failing to investigate and account for linguistic variability across various dimensions simultaneously. In addition, the potential effect of direction of language production on the affinity between translation and non-native language varieties has been neglected with only L1 translations (direct translation) involved. Directionality of the influence has also been incorporated into the first constraint dimension of language activation (Kotze, 2020: 346). Thus, L2 translations (inverse translation), as a bilingual language production, should be classified as one type of constrained language varieties. Otherwise, the conclusions regarding whether the constrained language universals exist or not may be biased. Given the scant study on constrained language from a typologically distant language pair and the existing methodological shortcomings, the present study deals with constrained language varieties of English translated from Chinese or written by native Chinese speakers. Dependency distance (DD), the linear distance between a word (dependent) and its head (governor) with syntactic dependency relations in a sentence (Liu, 2008), is commonly accepted as a reliable linguistic indicator of syntactic complexity and cognitive difficulty. DD has been widely applied

to studies on syntactic complexity in second language/interlanguage (Jiang et al., 2019), language processing mechanisms (Lu et al., 2016) and, more recently, translation studies (Fan & Jiang, 2019). Furthermore, the general tendency toward dependency distance minimization (DDM) has been found to be a linguistic universal related to the constraints of memory (Liu et al., 2017), and under this principle, the probability distribution of dependency distances (DD distribution) in human languages could fit well a long-tail power-law distribution, the right-truncated modified Zipf-Alekseev distribution (Ouyang & Jiang 2017), showing that the longer the DDs are, the less frequently they are used. Apart from the universality of human language reflected by general DD distribution, relevant parameters of the fitting formula may also reflect some linguistic peculiarities under different processing conditions. In this vein, two dependency-related indicators, mean dependency distance (MDD) and DD distribution are introduced to measure the syntactic similarities and differences among these linguistic varieties and to unveil their cognitive processing mechanisms. Constrained dependency treebanks with professional translated English (both L1 and L2) and non-native English and a reference treebank of non-constrained original English texts comparable in genre (short stories), time span (1930-2010) and text size (with a total of about 800,000 words) are set up.

This study attempts to trace the universal tendencies typical of constrained communication with Kruger's (2018) five constraint dimensions as an analytical framework. Two dimensions, modality/register (written/literary stories) and task expertise (expert) are controlled for and directionality of language production (both direct and inverse translations) is included to better compare and model the commonalities and divergences between varieties, and to disentangle the interplay of varying constraints on constrained language features. The following two research questions are addressed.

1) Do translated English and non-native English share similar syntactic features distinct from the native production in terms of mean dependency distance and parameters of probability distribution of dependency distance?

2) Does language directionality affect the affinity of translated English and non-native English regarding DD-related constrained syntactic features?

Results of the study show that translated English (both L1 and L2) and non-native English exhibit a significantly shorter MDD than non-constrained original English, whereas no significant difference is shown between the two constrained ones. This may imply that translators and non-native English writers are likely to produce syntactically less complicated sentences with a shorter MDD than the monolingual output due to shared cognitive constraints involved in bilingual activation. Furthermore, DD distribution of both constrained and non-constrained language varieties conform to the right truncated modified Zipf-Alekseev power-law distribution, yet a markedly minimized DD tendency, reflected by relevant parameters, prevails in constrained ones, again evidencing the bilingually constrained mechanism at work. Surprisingly, it seems that directionality, mediated more or less by other dimensions of constraints, does not play a big role as we expected in affecting the features of constrained languages, pending for further research. Integrating the length and frequency data, the dependency metrics of DD and parameters of its frequency distribution used in this study might provide novel

approaches to the syntactic linear characterization of constrained language varieties in a holistic, dynamic and systematic way, and thus shed light on the cognitive mechanism underlying bilingual production.

References

- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233-250). John Benjamins.
- Lanstyák, I., & Heltai, P. (2012). Universals in language contact and translation. *Across Languages and Cultures*, 13(1), 99-121.
- Kruger, H. (2018). Expanding the third code: Corpus-based studies of constrained communication and language mediation. In S. Granger, M.-A. Lefer, & L. Penha-Marion (Eds.), *Book of abstracts. Using corpora in contrastive and translation studies conference (5th edition)*. CECL papers 1 (pp. 9-12). Louvain-la-Neuve: Centre for English Corpus Linguistics/Université catholique de Louvain.
- Kruger, H., & van Rooy, B. (2016a). Constrained language: A multidimensional analysis of translated English and a non-native indigenised variety of English. *English World-Wide*, 37(1), 26-57.
- Granger, S., & Lefer, M.-A. (2020a). Introduction. In S. Granger, & M.-Aude Lefer (Eds.), *Translating and comparing languages: Corpus-based insights*. Corpora and language in use. *Proceedings* 6 (pp. 11-22). Louvain-la-Neuve: Presses universitaires de Louvain.
- Kotze, H. (2020). Converging what and how to find out why an outlook on empirical translation studies. In L. Vandevoorde, J. Daems, & B. Defrancq (Eds.), *New Empirical Perspectives on Translation and Interpreting* (pp.333-371). NY: Routledge.
- Gaspari, F., & Bernardini, S. (2010). Comparing non-native and translated language: Monolingual comparable corpora with a twist. In R. Xiao (Ed.) *Using corpora in contrastive and translation studies* (pp. 215-234). Newcastle: Cambridge Scholars Publishing.
- Rabinovich, E., Sergu, N., Noam O., & Shuly W. (2016). On the similarities between native, non-native and translated texts. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 1870-1881*. Berlin: ACL.
- Ferraresi, A. (2019). Collocations in Contact: Exploring Constrained Varieties of English through Corpora. *Textus*, (1), 203–222. doi: 10.7370/93190
- De Sutter, G., & Lefer, M.-A. (2020). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives*, 28(1), 1-23.
- Ivaska, I., & Bernardini, S. (2020). Constrained language use in Finnish: A corpus-driven approach. *Nordic Journal of Linguistics*, 43(1), 33–57.

- Kajzer-Wietrzny, M., & Ivaska, I. (2020). A Multivariate Approach to Lexical Diversity in Constrained Language. *Across Languages and Cultures*, 21(2), 169-194.
- Kajzer-Wietrzny, M., & Grabowski, Ł. (2021). Formulaicity in constrained communication: An intermodal approach. *MonTI. Monografías de Traducción e Interpretación*, 13, 148-183. doi: 10.6035/MonTI.2021.13.05
- Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, (2), 159-191.
- Jiang, J., Ouyang, J., & Liu, H. (2019). Interlanguage: A perspective of quantitative linguistic typology. *Language Sciences*, 74, 85-97.
- Lu, Q., Xu, C., & Liu, H. (2016). Can chunking reduce syntactic complexity of natural languages?. *Complexity*, 21(S2), 33-41.
- Fan, L., & Jiang, Y. (2019). Can dependency distance and direction be used to differentiate translational language from native language? *Lingua*, 224, 51-59.
- Liu, H., Xu, C., & Liang, J. (2017). Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*, 21, 171-193.
- Ouyang, J., & Jiang, J. (2018). Can the Probability Distribution of Dependency Distance Measure Language Proficiency of Second Language Learners? *Journal of Quantitative Linguistics*, 25(4), 295-313. doi: 10.1080/09296174.2017.1373991

Trados and the CRITT TPR-DB: Translation Process Research in an Ecologically Valid Environment

Longhui Zou, Michael Carl
Kent State University

Much of the translation process research (TPR) has been conducted with Translog-II, which is a research tool that makes it possible to record keystrokes and eye tracking data during translation sessions (Carl, 2012). The collected data can then be uploaded to the CRITT Translation Process Research Database (CRITT TPR-DB) which provides numerous tools for data analysis and data visualization (Carl et al., 2016). However, Translog-II is a research tool that does not offer editing possibilities, nor a translation environment to which professional translators are used. To increase the ecological validity of TPR, we have implemented a new interface that integrates keystroke and eye-tracking data collected by the widely used computer-assisted translation (CAT) tool, SDL Trados Studio. The data collected in Trados is synchronized with the eye tracker data and converted into the data format used by the CRITT TPR-DB.

SDL Trados Studio has a non-invasive plugin called *Qualityity* that runs in the background and collects the keystroke activities from the translator¹. Qualityity captures each keystroke and assigns it a time stamp as well as the segment number in which the translator is working. These data can be converted into a Translog-compatible XML format and integrated into the CRITT TPR-DB via the newly added upload option in the CRITT TPR-DB management tool².

However, unlike Translog-II, Qualityity does not offer the possibility to connect directly to external eye trackers. Qualityity also does not record where on the screen (X/Y positions) a word was edited or a segment occurs. Nevertheless, in order to investigate translator's gazing behavior during the translation process, we designed an add-on to the Trados-Translog-II conversion tool that integrates eye tracking data with the Qualityity keystroke data in a seamless way. This allows us to exploit user activity data (UAD) collected during translation sessions in Trados as a combination of eye movement and keyboard logging. Due to the different ways in which the gaze data is recorded in Translog-II and within the Trados setting some different processing strategies are mandatory and have been put in place.

Trados-Qualityity and the eye tracker (Tobii, Eyelink, GazePoint³) run as two independent programs, each of which is equipped with an independent keylogger. That is, every keystroke pressed by the translator in Trados is logged twice, once by Qualityity, and once by the eye tracker software, each with independent timers and timestamps. However, both trackers are configured such that the time is recorded by millisecond, which makes the two streams of events

¹ See <https://appstore.sdl.com/language/app/qualityity/612/>

² <https://critt.as.kent.edu/cgi-bin/yawat/yawat.cgi>

³ Tobii and Eyelink provide tools to extract tables with gaze and keystroke data. For GazePoint we implemented our own tool that runs a keylogger.

comparable. Finding the right offset between the timestamps recorded by Qualityity and within the eyetracker output allows us to synchronize gaze data with Qualityity keystrokes. While the eye tracker provides us with a stream of X/Y coordinates that reveal where on the screen the translator is looking at a certain point in time, Qualityity tells us for each keystroke the segment that is edited at a given point in time. The synchronized information allows us to relate the course of gaze events (X/Y coordinates) with (sequences of) keystrokes that occur in a certain Trados segment. In addition, while processing the data within the TPR-DB, the source and the target texts are aligned on a word-level, and each keystroke is associated with the (target) word that is produced. Thus, just as with Translog-II, the collected gaze data with the Trados setting can be synchronized with keystrokes and be mapped on source segments, via the target-source alignment. However, in contrast to Translog-II, we do not know which word(s) the translator is looking at. Whereas Translog-II computes gaze-to-word mapping at run-time, in the Trados setting we only know the X/Y coordinates of the gaze path and whether the gaze occurred in the ST or the TT windows. This information is, however, highly informative and reveals indicators of translation effort, as encoded, for instance, in the relative and total distances between fixations, number of regressions, gaze movements in the X and Y direction of the screen, parallel (concurrent) reading and typing behavior etc. Within the TPR-DB we attribute this gaze path summary information to various process and product units. For instance, gaze path information can be computed on the segment level (SG) for production units (PU), alignment groups (AG), or also for each word (ST), which provides novel ways to assess translation effort and effects on various levels of granularity. A small-scale feasibility study of this new research tool was done in Aachen, Germany (Heilmann, 2021). Three participants conducted from-scratch translation and two participants conducted post-editing of machine translation (PEMT). All the participants were professional translators with L1 in German and L2 in English. The two English STs of the experiment were extracted from popular science news texts, with 167 words and 186 words, respectively; the two German MTs of the experiment were extracted from Google Translate with 182 words and 184 words, respectively. We gained altogether 37 segments of TTs in this study. Each segment is characterized by a complex set of translation behavior.

In this pilot study, we were using Qualityity plugin for Trados Studio, version 2019, and Tobii TX300 eye tracker with its software Tobii Studio, version 3.3.2⁴. While the keylogging data collected by Qualityity can benefit our analysis on temporal and technical efforts conducted by the translator, the eye tracking data collected by Tobii studio can provide information about translator's cognitive effort during the experiment (Krings, 2001, Snover et al. 2006; Lacruz, 2017; Moorkens, 2018; Toral et al., 2018; Jia et al., 2019). We found that the new Trados-Translog interface can successfully be used to synchronize keystroke and gaze data from text production sessions into various data tables at different levels

⁴ See manual and specifications of Tobii 300X at <https://www.tobiipro.com/product-listing/tobii-pro-tx300/>; Tobii studio: <https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/setup/installing-tobii-studio/>.

of granularity, including the text, the segment, the alignment group, and TU. Our results of gaze-path analysis show the idiosyncrasies in different translators' behavior patterns and post-edited segments are clearly separated from from-scratch translation by the behavior of TT focused revision.

This tool enables translation scholars to monitor and compare translation behavior from translator's cognitive activity, to investigate the distribution of reading and typing activities from the beginning to the end of a translation process, and to have better understanding of factors that impact translation productivity. The collected and processed data can provide us with detailed information about effort and effect in the ecologically realistic translation experiments. We are also able to expand the features with customized algorithms or open-source NLP packages to conduct further analysis regarding the translation effect, such as automatic translation assessment tools (e.g., BLEU and COMET), and linguistic complexity metrics (e.g., LingX⁵). Several ongoing studies are using this new tool and more results are likely to emerge soon.

References

- Carl, M., Schaeffer, M., & Bangalore, S. (2016). The CRITT translation process research database. In *New directions in empirical translation process research* (pp. 13-54). Springer, Cham.
- Carl, M. (2012, May). Translog-II: a Program for Recording User Activity Data for Empirical Reading and Writing Research. In *LREC* (Vol. 12, pp. 4108-4112).
- Heilmann, Arndt, (2021) "Translator Activity during Computer Assisted Translation." in Book of Abstracts of the Applied Linguistics and Professional Practice Conference, page 15.
- Jia, Y., Carl, M., & Wang, X. (2019). How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study. *The Journal of Specialised Translation*, 31, 60-86.
- Krings, H. P. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes* (Vol. 5). Kent State University Press.
- Lacruz, I. (2017). 21 Cognitive Effort in Translation, Editing, and Post-editing. *The handbook of translation and cognition*, 386.
- Moorkens, J. (2018). Eye tracking as a measure of cognitive effort for post-editing of machine translation. *Eye tracking and multidisciplinary studies on translation*, 143, 55.
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers* (pp. 223-231).
- Toral, A., Wieling, M., & Way, A. (2018). Post-editing effort of a novel with statistical and neural machine translation. *Frontiers in Digital Humanities*, 5, 9.
- Walker, C., & Federici, F. M. (Eds.). (2018). *Eye tracking and multidisciplinary studies on translation* (Vol. 143). John Benjamins Publishing Company.

5 See <https://github.com/ContentSide/lingx>.

Note the Editor: A quantitative and qualitative study of editorial impact on literary translators' habitus and practice⁶

Noa Marom Tsadok
Bar Ilan University

While nowadays the name of the translator of a novel appears on the first page of the book, the translation editor's name often does not, and even when it does, it tends to be almost unnoticeable. This has lead readers, critics, and translation scholars to downplay, or even ignore, the impact the editor may have on the translated novel. However, the role played by the translation editor may have great implications for the translation process and product, and for the professional status of the translator – and therefore deserves a more central place in the contemporary study of translation. This talk takes a step towards filling the empirical gap in our knowledge of the role played by the literary translation editor in both the field of literary translation and in setting literary translational norms.

The proposed contribution is based on a doctoral thesis, which is a first attempt to investigate the working relationship between literary translators and literary translation editors in Israel, and quantitatively tests the impact of the latter on published texts, by methods of corpus analysis.

The professional relationship between literary translators and editors was investigated using a questionnaire distributed through translators' and editors' professional online forums, and by conducting personal interviews with 12 translators and 6 editors. The interviews and questionnaires were analyzed using concepts from the sociology of translation, which derives from Pierre Bourdieu's (1983) theory of habitus and field. The qualitative analysis of the interviews and questionnaires examines how the Cultural Capital of translators and editors affects their behavior and impact in the field of literary translation, and whether and how the habitus of literary translators is influenced by their work relationship with editors.

The study further sets to investigate the impact of editors on the creation and preservation of translational norms (Toury, 1995), particularly those of the translated literary dialogue. While early norms of translating literary dialogues into Hebrew have been identified in previous studies carried by Ben-Shahar (1998) and Weissbrod (1991), these have apparently undergone significant changes in the past 30 years, and the part that editors may have played in this process is hereby considered.

The quantitative aspect of the study included a corpus analysis of editorial changes in translated literary texts. The texts were provided by the translators and editors who interviewed for the study who were also willing to share their edited documents. The corpus includes 15 novels of the same sub-system (Even-Zohar, 2005): the novels are all considered non-canonical literature (i.e. not the classics) from the genres of suspense, thrillers, romance etc. This sub-system was chosen in order to focus on the specific supposed characteristics of the translators who

⁶The paper is based on a doctoral thesis held under the supervision of Dr. Galia Hirsch and Prof. Rachel Weissbrod, Department of Translation and Translation Studies, Bar-Ilan University.

produce these translations – a group of often uncertified professionals, who paradoxically achieved the relatively high-standing status of literary translators (and the Symbolic Capital it entails), yet often endure employment insecurity and stagnated translation rates (Sela-Sheffy, 2016).

So far, the qualitative findings have shown that publishing houses do not adhere to a set regulation (and often to any regulation at all) controlling work procedures between translators and editors. As a result, translators are often not requested to review changes implemented by editors in their translation, let alone approve them. The qualitative analysis suggests that this has an impact on the professional habitus of literary translators, as it often affects their professional self-esteem, professional choices, and expectations from employers later on.

The corpus analysis has shown that translation editors have a substantial stylistic impact on the final version of the text, as they make a large number of stylistic changes, as opposed to ones that correct unequivocally grammatical or technical mistakes. Furthermore, while input provided by interviewees has shown that both translators and editors hold similar views and are guided by the same norms when treating literary dialogue – such as using certain colloquial and even "subversive" forms – data analysis has demonstrated that editors follow these norms more freely. This appears to be related to the editor's professional position, which is that of the resident language expert, and to a sense of confidence that any unconventional choice made by the editor would be viewed as conscious and deliberate, whereas translators might worry that they are merely viewed as a dime-a-dozen service providers. Translation fees and the professional insecurity of freelance life also affect translators' said attitude: they might worry that unconventional translational choices would be too marked, and would result in the publishing house severing its ties with them (Weissbrod, 2007).

The findings support the notion that when discussing the sociological aspects of literary translation and the agents that produce it, it is important to consider the crucial, yet understudied, role of the translation editor. The qualitative findings suggest that the editing stage has the potential to influence translators' habitus, self-esteem, and professional aspects, and the quantitative findings demonstrate the editor's influence on the literary translation process and product, to the extent of playing a major role in the creation and preservation of translational norms.

References

- Ben-Shahar, R. (1998). "The Language of Plays Translated into Hebrew from English and French - A Cultural-Stylistic Study." *Meta: journal des traducteurs/Meta: Translators' Journal*, 43(1), 54-67.
- Bourdieu, P. (1983). "The field of cultural production, or: The economic world reversed." *Poetics*, 12 (4-5), 311-356.
- Even-Zohar, I. (2005). Polysystem theory (revised). *Papers in culture research*, 38-49.
- Sela-Sheffy, R. (2016). "Elite and non-elite translator manpower: the non-professionalized culture in the translation field in Israel." *The Journal of Specialized Translation* 25, 54-73.

- Toury, G. (1995). *Descriptive translation studies and beyond*. Amsterdam: John Benjamins.
- Weissbrod, R. (1991). "Translation of Prose-Fiction from English to Hebrew: A Function of Norms (1960s and 1970s)", in: *ATA (American Translators' Association)*, vol. V, Mildred L. Larson (ed.). New York: State University of New York, pp. 206-223.

Literalization in the self-revision process of novice and experienced biomedical translators

Susana Valdez
Leiden University

Translation is assumed to be governed by norms shared within a specific community. These norms inform what is considered appropriate and inappropriate behaviour in a particular context (Toury, 2012, p. 63). A translator's success is often linked to her competence to navigate through existing alternative and competing norms, knowing what norms are applicable and not applicable to a particular context (e.g., target culture, language and text-type), and understanding when, how, and why such norms should be applied.

Descriptive studies of translation norms do not usually adopt process-oriented methods to investigate the underlying decision-making processes that occur during translation. However, adopting these approaches can provide a fresh perspective into how translators negotiate alternative and sometimes conflicting norms when drafting their final versions. In this study, process-oriented methods are used to capture the unfolding of the translation process, primarily focusing on what happens when these translators write and rewrite textual solutions to problematic translation units during self-revision.

Self-revision is a commonly-used notion in translation, in both academia and the industry, and yet it is difficult to define precisely what happens during self-revision. Commonly assumed to be one of the essential steps in the translation process, few studies have investigated self-revision systematically combining product- and process-oriented studies in medical or biomedical translation (see Alves and Vale 2011 for an exception). When it comes to self-revision, “[d]ifferent people do the job quite differently”, as observed by Mossop (2019, p. 191). Self-revision may take place at an identifiable, distinct moment, sometimes even hours or days after the translation phase. However, these are not the processes that this study investigates. This study looks at self-revision as a text production process that is not “confined to a separate or redrafting phase” (Shih, 2007, p. 296) since what translation observation has shown us is that some translators translate-by-revising (Mossop, 2019, p. 192) and in their case it is not possible or even desirable to differentiate between the drafting phase and the self-revision phase.

This process provides an excellent opportunity to study decision-making, which is precisely what this study is interested in. In other words, the unfolding of translation solutions from interim versions to the final version. These interim versions, the intermediate solutions discarded by the translator during the decision-making process that do not surface in the target text, are rarely accessible to a researcher analysing the product. Yet by comparing the interim version with the final version, a researcher can examine and map the move from a more literal version to a less literal one, or vice-versa, and better understand “the constraints to which translators

choose to subject themselves, and of the interdependencies and the relative force of them as constraints on the act" (Toury, 2012, p. 218).

Considering the above and adopting a process-oriented approach, this paper reports on a study that examines the self-revision process of novice and experienced translators of biomedical content in the pair English to European Portuguese. To collect the data, an experiment was designed to study thirty translations of a 244-word instructional text about a medical device intended for health professionals. The data elicited from fifteen novice translators with up to two years of experience in translation and fifteen experienced translators with eleven to twenty-nine years of experience included keylogging and screen-recording data, interim versions, and target texts. These data were analysed and triangulated to describe the translation solutions in the interim and final versions in response to problematic translation units. In other words, to test the hypothesis that novice and experienced translators tend to move from more literal versions to less literal ones in biomedical translation during the self-revision process. Following on from the work of Englund Dimitrova (2005), translation processed of novice translators will be compared to those of experienced translators to further understand if the tendency to move from more literal versions to less literal ones is less predominant in novice translators, or vice-versa.

Interim solutions have been suggested to be a source of valuable information to empirically examine Ivir's (1981, p. 58) hypothesis that translators proceed from identical-meaning formal correspondences to not-quite-identical meanings or to structural and semantic shifts only when the first are not available (Toury 2012, p. 225–26). This, in turn, led to Chesterman's literal translation hypothesis, defined as the tendency to proceed from more literal versions to less literal ones (Chesterman 2011, p. 26). The assumption is that the cognitive process tends to be influenced in the first stages by the formal features of the source text (see Halverson 2015).

To identify the problematic translation units a nuanced classification based on primary and secondary indicators of translation problems is proposed (building on Krings 1986; and Göpferich 2010). These indicators allow for a distinction between (a) non-problematic units and problematic units, and (b) interim solutions and consciously postponed decisions. After the problematic units of translation are identified, these are matched with the translation solutions found in the interim versions and target texts. The translation solutions are then classified adopting and adapting Chesterman's (2016) proposal of syntactic, semantic and pragmatic translation solution types. These solution types are further interpreted as source- or target-oriented.

Contrary to expectations, the analysis points towards a literalisation phenomenon in the translators' processes. The data also indicates that the tendency to proceed from less literal versions to more literal ones is more pronounced in novice translators than experienced translators. The findings reported here shed new light on the self-revision processes of novice and experienced translators and their relationship with prevailing translation norms and enable us to better understand the practices in place in professional biomedical translation.

References

- Alves, F., & Vale, D. C. (2011). On drafting and revision in translation: a corpus linguistics oriented analysis of translation process data. *Translation: Corpora, Computation, Cognition*, 1(1), 105–122. <https://doi.org/10.5281/zenodo.283500>
- Chesterman, A. 2011. "Reflections on the Literal Translation Hypothesis." In *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*, edited by Cecilia Alvstad, Adelina Hild, and Elisabet Tiselius, 23–35. Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/btl.94.05che>.
- . 2016. *Memes of Translation*. Revised. Amsterdam and Philadelphia: John Benjamins. <https://doi.org/10.1075/btl.123>.
- Englund Dimitrova, B. (2005). *Expertise and Explicitation in the Translation Process*. John Benjamins.
- Göpferich, S. 2010. "Data Documentation and Data Accessibility in Translation Process Research." *The Translator* 16 (1): 93–124. <https://doi.org/10.1080/13556509.2010.10799295>.
- Halverson, S. L. 2015. "Cognitive Translation Studies and the Merging of Empirical Paradigms." *Translation Spaces. A Multidisciplinary, Multimedia, and Multilingual Journal of Translation* 4 (2): 310–40. <https://doi.org/10.1075/ts.4.2.07hal>.
- Ivir, V. 1981. "Formal Correspondence vs. Translation Equivalence Revisited." In *Theory of Translation and Intercultural Relations [Poetics Today 2:4]*, edited by Itamar Even-Zohar and Gideon Toury, 51–59. Tel Aviv: Porter Institute for Poetics and Semiotics, Tel Aviv University.
- Krings, H. 1986. *Was in Den Köpfen von Übersetzern Vorgeht*. Tübingen: Gunter Narr.
- Mossop, B. (2019). *Revising and Editing for Translators*. Routledge. <https://doi.org/10.4324/9781315158990>
- Shih, C. Y. (2007). Revision from translators' point of view: An interview study. *Target*, 18(2), 295–312. <https://doi.org/10.1075/target.18.2.05shi>
- Toury, G. 2012. *Descriptive Translation Studies – and Beyond*. Amsterdam: John Benjamins. <https://doi.org/10.1075/btl.100>.

Using semi-structured interviews to revisiting the decision tree model for post-editing tasks: What can the language industry teach us?

Jean Nitzke², Carmen Canfora¹, Silvia Hansen-Schirra¹

¹University of Mainz

²University of Agder

Post-editing (PE) machine translation has become increasingly important in the professional translation industry in recent years. However, not every translation job is automatically suitable for machine translation (MT) and different aspects have to be taken into account to decide for or against PE. In fact, there is even a continuum of options: no PE, light PE, full PE, full PE plus revision or human translation, i.e. translation from scratch, with or without further quality assurance measures like revision or in-country review. We recently published a decision tree for post-editing jobs that shall guide decision makers in deciding whether a job is suitable for post-editing and if so what kind of quality assurance might lead to fit-for-purpose translations. The different branches of the decision tree cover aspects such as the availability of an MT system, characteristics of the source text (type), specificities of the target culture, time pressure, budget, etc.

Our approach is based on a risk and process analyses, which allows predictions on the potentials but also on the risks MT and PE may cause. This evaluation of opportunities and risks concerns both strategic decisions, e.g. whether machine translation should be used at all, and operational decisions. Operational decisions include assessing which text types or which specific projects involve high enough risks for the end users of the translation or the client that machine translation output by itself would be too risky and further quality assurance measures are necessary. Decisions for or against quality assurance measures such as light PE or full PE should therefore depend to a large extent on the risk level of a translation. Another parameter that plays a role in the risk assessment of the planned translation and thus ultimately in the decision-making process on appropriate quality assurance is the fundamental need for security of the respective client. This need for security can include both the avoidance of financial losses, for example due to reputational risks, and injuries, for example due to errors in safety-critical documents. Within the framework of a risk management approach, such as the one described in the relevant international standard ISO 31000 (ISO 2018), this need for security can be addressed by providing information about translation risks and thus improving risk awareness.

Our rather theoretical assumptions underlying the model present an idealised approach to post-editing projects and the market reality may diverge from our expectations. Therefore, our research aims at comparing empirical evidence from market stakeholders with our existing model and assess the potentials to update our model, because we believe theoretical models should reflect and be fed from realistic processes and environments, even though processes established in everyday practise may not be prototypical.

Hence, we developed a semi-structured interview building on our theoretical approach addressing language service providers, project managers, and institution that work with post-editing projects and have to make the decisions dealt with in our

model. The interview consists of 20 open questions and a ranking task, which will allow us to combine both qualitative and quantitative analyses. The questions cover aspects like the role of text types, risk awareness and risk assessment, target recipients, final quality requirements, money and time frame and more aspects conceptualising the post-editing project. We also ask about their attitude towards and the availability of machine translation systems as well as their experiences with guidelines.

Based on approximately 20 interviews, we will describe and rank which of the assumed aspects that influence the PE decision making process are indeed important in real-life projects, how they intertwine, and how risk management considerations as well as liability issues affect these decisions. The findings will provide innovative insights on how the translation service industry deals with PE projects, whether there are similar patterns between different actors, and, finally, how to adjust our model. In this scope, we have to discuss potential mismatches between an idealised model and the reality established at the translation market, e.g. the differences between what role the budget for a translation/post-editing project should play ideally and does play in real projects.

In the presentation, we will discuss the existing decision tree model, its flaws, the interview scheme that resulted from the flaws, and the analysis pattern we developed for this study. Furthermore, we will present anonymised information on our participants taken from the meta data collected within our study and findings concerning general MT use, post-editing processes, risk awareness, market pressures concerning time and price constraints and their generalised decision-making procedures. Finally, we will feed back these results into the existing model and present the optimised decision tree for post-editing tasks. Further, we will be discussing implications for curricula contents both for post-editing and project management competences.

References

ISO (International Standardization Organization). 2018. *ISO 31000:2018 Risk management – Guidelines*. Geneva: ISO.

Explicitation and implicitation in L1>L2 student translation: A relevance-theoretic approach

Laura Aguiar
UCLouvain

'Explicitation' is one of the constructs that has received most scholarly attention in corpus-based translation studies (Granger & Lefer, in press, p. 27). It has served as a starting point for a multitude of empirical works (see, e.g., Baumgarten, Meyer, & Özçetin, 2008, for an overview of the methodological approaches taken), with substantial supporting evidence being produced. For instance, translated English has been found to be significantly more explicit than non-translated English in terms of encoding of optional that-complementation in different registers (e.g., Kruger & Van Rooy, 2012; Olohan & Baker, 2000; Redelinguhuys & Kruger, 2015).

Despite such interest, however, the conceptualization of the construct has often been dealt with rather intuitively, without much theoretical elaboration (Krüger, 2013, p. 286; see also Krüger, 2014). In an attempt to elucidate this highly complex construct, notably in relation to its counterpart, 'implicitation,' which has attracted less attention, and establish a theoretical framework for their study, some researchers have turned to linguistic theories. Thus, Hansen-Schirra, Neumann, and Steiner (2007) refer to semantics and pragmatics to disentangle the concept of explicitness (i.e., a linguistic property of translated texts related to their internal, intratextual configuration) from that of explicitation proper (i.e., a linguistic property of translated texts resulting from an intertextual relationship, namely that between source and target texts). Krüger (2013) resorts to cognitive linguistics in general, and Langacker's model of construal operations and semantic theory of domains in particular, to model explicitation and implicitation. Murtisari (2013) and De Metsenaere and Vandepitte (2017) do so seeking recourse in relevance theory — a cognitive-pragmatic theory of communication developed by Sperber and Wilson (1986/1995).

Drawing on De Metsenaere and Vandepitte (2017), the present study seeks to illustrate how relevance theory can help define and empirically investigate the concepts of explicitation and implicitation, while discussing practical difficulties involved in the application of relevance-theoretic notions (e.g., context) to translation. In the study, explicitation and implicitation are approached from a unique perspective: that of student translations produced from the foreign/second language (L2) into the native/first language (L1) (L2>L1 translation) and the other way around (L1>L2 translation). The study aims to answer the following research questions: Are explicitation and implicitation patterns in L1>L2 student translations influenced by translation directionality? If so, how?

The research questions are addressed using corpus data extracted from the Multilingual Student Translation (MUST) corpus (Granger & Lefer, 2020). The data consist of 44 English>French translations (15,826 tokens) and 44 French>English translations (12,362 tokens), produced by the same 43 students enrolled in a Master's program in translation. All the students have French as L1, English as L2 (average of 9.7 years of study), and varying levels of experience in L1>L2

translation (the majority has undergone more training in L2>L1 translation than in L1>L2 translation). The translations derive from four English and four French source texts (viz., informative news articles), each containing about 300 words. In light of the new research agenda for empirical translation studies (De Sutter & Lefer, 2019), which promotes multifactorial analyses, factors (other than directionality) likely to affect explicitation and impication patterns in translations are taken into account. From a constrained communication standpoint (see, e.g., Kotze, 2020; Kruger & Van Rooy, 2016, for an outline of this theoretical framework), those include L2 English proficiency (as learners of English, the translators in question arguably display lower levels of production proficiency), translation experience (having received less training in L1>L2 translation, they also presumably have less translation experience in this direction), and potential source text and student variability.

The data are analyzed from a bilingual bidirectional parallel approach (i.e., comparing translations in French and English with their corresponding source texts in English and French) and manually annotated for explicitation and impication shifts in a text editor adopting an XML-inspired system. Building on relevance theory and on De Metsenaere and Vandepitte (2017, p. 402), explicitation is defined as the result of a process whereby propositional content that is not articulated in the source text utterance, but that can be recovered from an explicature (i.e., developed or inferred locally, from a subpropositional constituent) or implicature (i.e., inferred globally, from a fully propositional form) associated with the source text utterance, is linguistically encoded in the target text utterance. Impication is defined as the result of a process whereby propositional content linguistically encoded in the source text utterance is left unexpressed in the target text utterance, being however recoverable from an explicature or implicature associated with the target text utterance. Relevant explicatures and implicatures (i.e., triggered by specific target text utterances and helpful in identifying explicitation and impication shifts) are formulated for source text utterances. A total of 904 explicitation and 307 impication shifts are annotated.

Statistical testing is done in R (R Core Team, 2020) by means of mixed-effect modeling (lmer), using: (i) translation direction (L2>L1 vs. L1>L2), years studying English (viz., a proxy for L2 English proficiency), and translation experience (1 [limited experience in L1><L2 translation] vs. 2 [extensive experience in L1><L2 translation]) as explanatory variables; (ii) the absolute frequency of explicitation and impication shifts as response variables; (iii) source texts' alphabetical ID and students' numerical ID as random effect terms. One model is run per response variable. Preliminary results indicate that L2 English proficiency appears to be a better predictor for explicitation patterns ($p < 0.05$) than translation direction ($p = 0.18$) and translation experience ($p = 0.85$), with more proficient students explicating source texts to a greater extent, and that part of explicitation variation can also be attributed to both source texts (97.5% CI; [0.23 – 0.92]) and students' idiosyncrasies (97.5% CI; [0.22 – 0.64]). Additionally, they suggest that impication patterns can be explained by translation direction ($p < 0.001$) and L2 English proficiency ($p < 0.001$). Students in L2>L1 translation settings and with higher levels of L2 English proficiency impicate source texts to a greater degree. The present study hopes to contribute toward the theorization of explicitation and impication,

while showcasing how relevance theory can be useful in investigating a variety of explicitation- and implicitation-related linguistic phenomena.

References

- Baumgarten, N., Meyer, B., & Özçetin, D. (2008). Explicitness in translation and interpreting: A critical review and some empirical evidence (of an elusive concept). *Across Languages and Cultures*, 9(2), 177–203. doi:10.1556/Acr.9.2008.2.2
- De Metsenaere, H., & Vandepitte, S. (2017). Towards a theoretical foundation for explicitation and implicitation. *Trans-kom*, 10(3), 385–419.
- De Sutter, G., & Lefer, M.-A. (2019). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives: Studies in Translation Theory and Practice*, 28(1), 1–23. doi:10.1080/0907676X.2019.1611891
- Granger, S., & Lefer, M.-A. (2020). The Multilingual Student Translation corpus: A resource for translation teaching and research. *Language Resources and Evaluation*, 1–17. doi:10.1007/s10579-020-09485-6
- Granger, S., & Lefer, M.-A. (in press). Corpus-based translation and interpreting studies: A forward-looking review. In S. Granger & M.-A. Lefer (Eds.), *Extending the scope of corpus-based translation studies* (pp. 13–41). London, England: Bloomsbury.
- Hansen-Schirra, S., Neumann, S., & Steiner, E. (2007). Cohesive explicitness and explicitation in an English-German translation corpus. *Languages in Contrast*, 7(2), 241–265. doi:10.1075/lic.7.2.09han
- Kotze, H. (2020). Converging what and how to find out why: An outlook on empirical translation studies. In L. Vandevoorde, J. Daems, & B. Defranq (Eds.), *New empirical perspectives on translation and interpreting* (pp. 333–370). Abingdon, England: Routledge.
- Kruger, H., & Van Rooy, B. (2012). Register and the features of translated language. *Across Languages and Cultures*, 13(1), 33–65. doi:10.1556/Acr.13.2012.1.3
- Kruger, H., & Van Rooy, B. (2016). Constrained language: A multidimensional analysis of translated English and a non-native indigenised variety of English. *English World-Wide*, 37(1), 26–57. doi:10.1075/eww.37.1.02kru
- Krüger, R. (2013). A cognitive linguistic perspective on explicitation and implicitation in scientific and technical translation. *Trans-kom*, 6(2), 285–314. Retrieved from <http://www.trans-kom.eu>
- Krüger, R. (2014). From S-explicitation to T-explicitation? Tracing the development of the explicitation concept. *Across Languages and Cultures*, 15(2), 153–175. doi:10.1556/Acr.15.2014.2.1
- Murtisari, E. T. (2013). A relevance-based framework for explicitation and implicitation in translation: An alternative typology. *Trans-kom*, 6(2), 315–344. Retrieved from <http://www.trans-kom.eu/>
- Olohan, M., & Baker, M. (2000). Reporting that in translated English: Evidence for subconscious processes of explicitation? *Across Languages and Cultures*, 1(2), 141–158. doi:10.1556/Acr.1.2000.2.1

- R Core Team. (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>
- Redelinghuys, K., & Kruger, H. (2015). Using the features of translated language to investigate translation expertise: A corpus-based study. *International Journal of Corpus Linguistics*, 20(3), 293–325. doi:10.1075/ijcl.20.3.02red
- Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Hoboken, NJ: Blackwell Publishing.

Register Variation in Student Translations: A Study on Impersonal Verb Use in Multiple Translation Corpora

Tiffany Jandrain
Université de Mons

Register variation has increasingly been observed in Translation Studies (Baker, 1992/2011; Granger, 2016; Hansen-Schirra et al., 2012; Lapshinova-Koltunski, 2013, 2014; Lefer & Voegele, 2016; Neumann, 2012, 2016, 2021; Schaffner, 2002). In fact, the importance of analyzing registers before translating texts has been recently emphasized (Gledhill & Kubler, 2016). As Neumann (2021) indicated, “if they [translators] are not aware of register conventions or deem it fit to depart from them, we might be able to observe dilution of register features in translated texts” (p. 66).

This paper focuses on register analysis in translations made by Master students translating from English into French. Studies have shown that students may find register transposition difficult when they translate non-literary texts because of the differences between English and French registers (Chuquet & Paillard, 1987/1989; Gadet, 1996; Vinay & Darbelnet, 1958). Even if students may often feel that language needs to be adapted according to register variables (such as the audience or the function of the target text), many of them fail to do this successfully (Fawcett, 1997; Gile, 2005; Vandaele, 2015).

Our analysis was conducted on impersonal verbs (IVs) in French, which modify the message organization. In fact, they “rearrange” the communication (Riegel et al., 2018, p. 666), since the real subject is postponed (Bottineau, 2010) and does not introduce the theme of the sentence, which leads to its depersonalization. IVs (such as “*il y a*” (“there is”) or “*il semble que*” (“it seems that”)) may also be a stylistic marker, a genre-specific feature, since, regarding academic texts, “impersonality [i.e. the use of formal, detached language, avoiding reference to the author] is seen as a defining feature of expository writing as it embodies the positivist assumption that academic research is purely empirical and objective” (Hyland, 2002, p. 1095, in Cigankova, 2016, p. 100). Therefore, IVs seem to depend on the language variety in which they are used, or “register” as defined by Biber and Conrad (2009), to whom a register is “language variety associated with both a particular situation of use and with pervasive linguistic features that serve important functions within that situation of use” (p. 31).

This paper shows the results of a study carried out on the students’ use of IVs in translations in comparison with the use of IVs in non-translated texts. In fact, a variation between natural language (of non-translated texts) and translated language can be hypothesized (Zanettin, 2013, in Kruger, 2018). The two research questions of this paper are the following:

- 1) Is the use of IVs in French by translation students different from the authors’ use in nontranslated texts?
- 2) To which extent does register have an influence on the use of IVs in natural and translated language in French?

A corpus-based study was conducted following three steps: (1) an analysis of the use of IVs in nontranslated texts considered comparable to the students' translations; (2) an analysis of their use in the students' translations in comparison with their source texts; and (3) a comparison between the results of the first two steps. Since a register analysis implies the analysis of at least two registers (Neumann, 2021), two registers, that are taught to Master students at University of Mons, Belgium, were under scrutiny: press article and popular science article.

Four corpora were thus compiled: two comparable corpora (one per register) of 30 French nontranslated texts, and two corpora (one per register) of 14 translated texts written by students in French (i.e. translations of a press article from *The Economist* (TT1) and translations of a popular science article from *New Scientist* (TT2)). It should be noted that those translation corpora contain 14 translations of a same source text and can therefore be referred to "multiple translation corpora", which are of great interest since that type of corpus "allows for comparison of translation solutions [for a same linguistic unit] used by various students rendering the same text" (Granger & Lefer, 2020, p. 1186). IVs were extracted in our corpora by searching for the words "il" and "ca" (which are the subjects of IVs in French by definition) in the *Sketch Engine* concordancer (Kilgariff et al., 2014).

Results of the comparable corpora show that there is a significant difference in personal and impersonal verb use according to the registers of natural language ($\chi^2(1) = 14.865$, $p < 0.05$, Cramer's $V = 0.058$). IVs are more frequent in the popular science articles. Some IVs in press articles tend to imply readers in the message (e.g. "*il faut supposer*" ("we must suppose")), while IVs in popular science articles tend to provide readers with new information (e.g. "*il est question de*" ("it is a matter of")) or expose the scientists' points of view (e.g. "*il est prématuré*" ("it is too soon")). Besides, some IVs do not make texts impersonal because they express opinions (e.g. "*il est déplaisant*" ("it is unpleasant")).

Results from the multiple translation corpora show that there is a significant difference in personal and impersonal verb use according to the registers of translated language ($\chi^2(1) = 6.3364$, $p < 0.05$, Cramer's $V = 0.081$) as well. There was, however, no significant difference in the use of those verbs between the students' texts ($p > 0.05$ in TT1 and $p > 0.05$ in TT2). Results also show that 50% of the students used IVs in their TT1, while 100% of them decided to include IVs in their TT2. It should be noted that 75% of IVs in TT2 were literal translations, while 30% of IVs in TT1 were spontaneously added by the students. Finally, the analysis of each occurrence of IVs compared to the source text units shows that 40% of IVs in TT1 and 2,5% of IVs in TT2 led to meaning mistakes.

From a qualitative point of view, results of the comparison between non-translated and translated texts globally show that students tend to use IVs in a similar way than authors do in natural language. In our corpora, IVs seem more problematic to students when they translate press articles.

References

- Baker, M. (2011). *In other words: A coursebook on translation* (2nd ed.). Routledge. (Original work published 1992)
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge University Press.

- Bottineau, D. (2010). Quand le classement est une théorie: Le verbe impersonnel dans les verbes français. *Langages*, 179–180, 57–77. <https://doi.org/10.3917/lang.179.0057>
- Chuquet, H., & Paillard, M. (1989). *Approche linguistique des problèmes de traduction anglais-français* (2nd ed.). Ophrys. (Original work published 1987)
- Cigankova, N. (2016). Grammatical expression of impersonality in LSP texts and translations. *Procedia - Social and Behavioral Sciences*, 231, 99–106. <https://doi.org/10.1016/j.sbspro.2016.09.077>
- Fawcett, P. (1997). *Translation and language*. St. Jerome Publishing.
- Gadet, F. (1996). Niveaux de langue et variations intrinsèques. *Palimpsestes. Niveaux de Langue et Registres de la Traduction*, 10, 17–40. <https://doi.org/10.4000/palimpsestes.1504>
- Gile, D. (2005). *La traduction: La comprendre, l'apprendre*. Presses Universitaires de France.
- Gledhill, C., & Kubler, N. (2016). What can linguistic approaches bring to English for Specific Purposes? *ASp. Concepts and Frameworks in English for Specific Purposes*, 69, 65–95. <https://doi.org/10.4000/asp.4804>
- Granger, S. (2016). A lexical bundle approach to comparing languages. Stems in English and in French. In M.-A. Lefer & S. Vogeleer (Eds.), *Genre- and register-related discourse features in contrast* (pp. 59–72). John Benjamins Publishing Company.
- Granger, S., & Lefer, M.-A. (2020). The Multilingual Student Translation corpus: A resource for translation teaching and research. *Language Resources and Evaluation*, 54(4), 1183–1199. <https://doi.org/10.1007/s10579-020-09485-6>
- Hansen-Schirra, S., Neumann, S., & Steiner, E. (2012). *Cross-linguistic corpora for the study of translations. Insights from the language pair English-German*. De Gruyter Mouton.
- Kilgarrieff, A., Baisa, V., Bušta, J., Jakubiček, M., Kovař, V., Michelfeit, J., Rychly, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
- Kruger, H. (2018). *Expanding the third code: Corpus-based studies of constrained communication and language mediation*. Using Corpora in Contrastive and Translation Studies (5th ed.), Université Catholique de Louvain.
- Lapshinova-Koltunski, E. (2013). VARTRA: A comparable corpus for analysis of translation variation. *Proceedings of the Sixth Workshop on Building and Using Comparable Corpora*, 77–86. <https://aclanthology.org/W13-2510.pdf>
- Lapshinova-Koltunski, E. (2014). Variation in translation: Evidence from corpora. In C. Fantinuoli & F. Zanettin (Eds.), *New directions in corpus-based translation studies* (pp. 93–114). Language Science Press.
- Lefer, M.-A., & Vogeleer, S. (2016). Introduction. In M.-A. Lefer & S. Vogeleer (Eds.), *Genre- and Register-related Discourse Features in Contrast* (pp. 1–6). John Benjamins Publishing Company.
- Neumann, S. (2012). Register-induced properties of translations. In S. Hansen-Schirra, S. Neumann, & E. Steiner, *Cross-linguistic corpora for the study of translations. Insights from the language pair English-German* (pp. 191–209). De Gruyter Mouton.

- Neumann, S. (2016). Cross-linguistic register studies. In M.-A. Lefer & S. Vogeleeer (Eds.), *Genre- and register-related discourse features in contrast* (pp. 35–57). John Benjamins Publishing Company.
- Neumann, S. (2021). Register and translation. In K. Mira, J. Munday, W. Zhenhua, & W. Pin (Eds.), *Systemic Functional Linguistics and Translation Studies* (pp. 65–82). Bloomsbury Academic.
- Riegel, M., Pellat, J.-C., & Rioul, R. (2018). *Grammaire méthodique du français* (7th ed.). Presses Universitaires de Paris (puf). (Original work published 1994)
- Schaffner, C. (Ed.). (2002). *The role of discourse analysis for translation and in translator training*. Multilingual Matters LTD.
- Vandaele, S. (2015). La recherche traductologique dans les domaines de specialite: Un nouveau tournant. *Meta*, 60(2), 209–235. <https://doi.org/10.7202/1032855ar>
- Vinay, J.-P., & Darbelnet, J. (1958). *Stylistique comparée du français et de l'anglais*. Didier.

Metaphor awareness and metaphor translation competence: an empirical study on translation students

Charlene Meyers
UMONS

In recent years, pedagogical studies have turned to translation classes to analyze how translation students cope with metaphors in a source text (ST). Unsurprisingly, the pedagogical scope of translation studies has mainly focused on “metaphor translation competence,” which is defined as how translators translate metaphors from a source language to a target language (Sjorup, 2013). Moreover, some studies such as the one by Jensen (2005), have taken an interest in comparing the student’s level of metaphor translation proficiency with that of young professionals and experts. For instance, Jensen (2005) was able to demonstrate that students, when facing a metaphor in the ST, were less likely to choose a metaphorical solution in the target text than young professionals and experts.

However, aside from the ability of students to translate metaphors, little attention has been paid to their ability to identify and recognize a metaphor in a source or target language outside of a translation task, even though this parameter is said to directly influence their metaphor translation competence (Andersen, 2004). Andersen (2004) defines this ability as “metaphor competence,” which is the translator conscious knowledge or awareness of the metaphor as a pragmatic, linguistic, and cognitive item (according to the metaphor theory by Lakoff & Johnson, 1980).

The aim of the present study is to examine both the metaphor *translation* competence and metaphor competence of Master students in order to gauge their awareness of metaphors and evaluate their translation solutions for the metaphors in a ST to verify whether or not metaphor competence is an essential prerequisite for metaphor translation competence.

Metaphor translation competence was examined to provide an answer to the following research questions:

- Is the metaphoricity identified in the ST retained in the target text (TT) by translation students?
- Is there a difference in the proportion of metaphors among the TTs?
- Which translation strategies are applied to the metaphors? Is the frequency of strategies applied to the metaphors in the ST equivalent?

While metaphor competence was analyzed to answer the following research questions:

- Are translation students able to identify metaphors in a ST they had previously translated?
- Is the proportion of metaphors identified by translation students equivalent?
- Does the metaphoricity identified by translation students depend on the lexical unit?

To test metaphor translation competence, the translation of metaphors in a ST by 20 translation students was analyzed. A ST was selected in the *New Scientist* and was examined for metaphors by following the MIPVU (Metaphor Identification Procedure Vrije Universiteit Amsterdam) proposed by Steen et al. (2010). The metaphor-related words (MRWs) in the ST were coded by two researchers using the identification procedure on every lexical unit of the ST. Intercoder agreement was calculated on the identification of MRWs and showed, after three rounds, a Cohen's kappa of 0.958, which reveals an "almost perfect agreement" (McHugh, 2012, p. 279). In total, 75 MRWs were identified in the ST by both coders and were retained for the experiment.

20 Master students were then asked to sight translate this ST from English to French, as a regular sight translation exercise during their scientific and technical class. The recording of their sight translations was transcribed and the MIPVU (Steen et al., 2010) was applied by the same two researchers on the TTs to identify the MRWs. Intercoder agreement was also calculated on the TTs and showed "substantial agreement" ($0.61 < \kappa < 0.80$) for some TTs and "almost perfect agreement" ($0.81 < \kappa < 1.00$) for others.

The results show that the translation of MRWs in the ST is metaphorical in 80.26% of cases in the TTs and that there is no significant difference in the proportion of MRWs between TTs ($c2(19) = 25.186, p = 0.1545$). However, there is a significant difference in the proportion of translation strategies applied to individual MRWs ($c2 = 2744.8, p = 0.0004998$. Cramer's $V = 0.511$).

To evaluate metaphor competence, the ability of translation students to identify metaphors in the same ST was tested. 47 translation students received a two-hour seminar, which aimed to provide simplified explanations about metaphorology and the MIPVU coding method (Steen et al., 2010). Students were given a short manual for coding as well as examples and a spreadsheet with lexical units to code from the same ST they had previously translated. 21 MRWs were selected out of the 75 MRWs identified in the whole text based on the fact that they were evenly distributed in the text and also to avoid a long and tiresome coding experiment for students. Following the simplified explanations inspired from the MIPVU, students were asked to code the 21 lexical units that had all been identified as metaphorical by the researchers. More specifically, they were asked to apply either a code of 1 for "metaphorical items" or a code of 2 for "non-metaphorical items."

The results show that students were able to code items as metaphorical in 50.04% of the cases. They also show that there is a significant difference in the proportion of MRWs identified by students ($c2(46) = 63.951, p = 0.040$. Cramer's $V = 0.255$). Moreover, Krippendorff's alpha was computed to test the average coding agreement between students and resulted in a fair agreement ($\alpha = 0.282$). Finally, the results show that there is a significant association between the lexical units and the metaphoricity coded by students ($c2(20) = 292.87, p = 2.2 \cdot 10^{-16}$. Cramer's $V = 0.545$). In other words, the metaphoricity attributed by students is highly dependent on the lexical unit under scrutiny.

Interestingly, these results give rise to a paradox: although translation students have fairly good metaphor translation competence, their ability to identify MRWs and therefore their metaphor competence are not as good. In contrast with other

studies, this study stresses that metaphor competence is paradoxically not a prerequisite for metaphor *translation* competence.

References

- Andersen, M. S. (2004). Metaforkompetence—En empirisk undersøgelse af semi-professionelle oversætteres metaforviden [Metaphor competence—An empirical study of semi-professional translators' metaphor knowledge]. In *Copenhagen Working Papers in LSP*.
- Jensen, A. (2005). Coping with Metaphor: A cognitive approach to translating metaphor. *HERMES – Journal of Language and Communication in Business*, 18(35), 183–209.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/BM.2012.031>
- Sjorup, A. C. (2013). *Cognitive effort in metaphor translation: An eye-tracking and key-logging study* [Doctoral dissertation]. Copenhagen Business School.
- Steen, G., Dorst, A. G., Herrmann, J. B., Kaal, A., Krennmayr, T., & Trijntje, P. (2010). *A method for linguistic metaphor identification: From MIP to MIPVU*. John Benjamins.PVU. John Benjamins.

A new perspective on translation studies – translation as overcoming communication barriers

Sergio Hernández Garrido and Rebecca Schulz
University of Hildesheim

The prototypical understanding of *translation* has traditionally been the interlingual transfer between two languages in the sense of *translation proper* (Jakobson 1956, see also Hansen-Schirra/Maaß 2019). Recent developments in the translation market and new approaches and interests in research suggest that intralingual translation, which Jakobson defined as *rewording* or *paraphrase*, also falls into the concept of *translation*. Translating and interpreting into comprehensibility enhanced language varieties such as Easy Language, Easy Language Plus, and Plain Language (see Bredel/Maaß 2016, Hansen-Schirra/Maaß 2020, Maaß 2020) have become more relevant in the German translation market due to changes in the legal framework (see Lang 2020, 2021). It also has a predominant position in accessible communication research, where translation studies' approaches have been implemented for the development of rulebooks and translation strategies (see Bredel/Maaß 2016, Maaß 2020, Hansen-Schirra/Maaß/Rink 2021).

Functionalistic translation approaches such as Reiß/Vermeer (1984) already proposed that the language difference is not the only communication barrier relevant when translating. For instance, cultural barriers can also represent a challenge for target audiences. In this context, accessible communication research suggests that there are more communication barriers that have to be considered when creating content for specific target groups (see Schubert 2016, Rink 2019, 2020, Lang 2021, for English terminology see Maaß 2020). The basic assumption is that these target groups may face communication barriers if the content is not adapted to their communicative needs and expectations. A sensory barrier may come into place if the target groups cannot perceive the content. Texts can also represent a cognitive barrier if the information is too abstract for them to process. If the text's physical presentation is not suitable for the users, it represents a motor barrier. An expert knowledge barrier comes into question if domain specific knowledge is necessary to understand the text. This barrier is often accompanied by an expert language barrier if the language variety used is too domain specific. If the text's channel of distribution and media presentation do not correspond to the target groups' preferences and expectations, the text may represent a media barrier. And finally, a text can represent a motivational and emotional barrier, if the content does not appear useful or if it leads to strong emotions that consume their comprehension resources. Overcoming these communication barriers is perceived as the main goal of accessible communication experts, whose approaches often overlaps with the ones of a translator (see Bredel/Maaß 2016, Hansen-Schirra/Maaß 2019, Maaß 2020). This suggests that these barriers can be adopted into a new concept of *translation* where it is defined as overcoming communication barriers.

The steps towards accessibility can be obstructed by communication barriers (see Maaß/Rink 2019, Maaß 2020). Accessible and action-enabling content should have certain properties. First of all, they should be retrievable for the target groups; target

groups should be able to find the content with the means at their disposal. Accessible content should also be perceptible; target groups should perceive the content with the sensory channels at their disposal. Comprehensibility is also key for content to be accessible; target groups should be able to comprehend the text. If information is linkable to previous knowledge, it is easier for the target groups to recall information. Accessible content should also be acceptable for the target groups. And finally, content that is accessible should be action-enabling for the target groups; they should be able to act on the text's information. Each of these properties built on each. This accessibility concept supports translation professionals in the creation of content suited to their target groups communicative needs and expectations. It helps to recognize which steps are being obstructed by communication barriers in order to find strategies to overcome them.

In our presentation, we will present the communication barriers from a Translation Studies perspective. This approach can be seen as fundamental research which is the necessary basis for empirical projects. Recent studies split up into text and user perspective. Considering the text perspective, the focus is on linguistic features regarding word, sentence and text level as well as semantics and pragmatics (see Kröger in prep. and Hernandez-Garrido in prep.). This leads to inferences about text features of different text types which can challenge users with different communicative needs. On the other hand, research on the user perspective focusses on the user's global, grammatical and discourse knowledge (Maaß 2020: 92) which influences the access to information.

In recent studies like Keller (in prep.), Ahrens (in prep.) and Schulz (in prep.) the user perspective is empirically examined with different focal points. Both perspectives have to be considered separately from each other in a first step. Thus, the results can be combined in a second step in order to obtain results on the expedient design of accessible information.

References

- Ahrens, S. (in prep.): Patientinnenaufklärung beim Frauenarzt – Welche Eigenschaften müssen Einfache-Sprache-Texte haben, um für Frauen mit Deutsch als Zweitsprache verständlich und handlungsanleitend zu sein? Manuskript der Dissertationsschrift. Universität Hildesheim.
- Bredel, U./ Maaß, C. (2016): *Leichte Sprache: Theoretische Grundlagen, Orientierung für die Praxis*. Berlin: Dudenverlag.
- Hansen-Schirra, S./Maaß, C. (2019): Translation proper: Kommunikationsbarrieren abbauen. Uni Hildesheim. <https://doi.org/10.25528/015>
- Hansen-Schirra, S./ Maaß, C. (2020): *Easy Language Research: Text and User Perspectives*. Berlin: Frank & Timme
- Hernandez-Garrido, S. (in prep.): Enhancing the accessibility of health information: Audiovisual formats in the official COVID-19 communication in Colombia as a means to tackle health illiteracy. Manuskript der Dissertationsschrift. Universität Hildesheim.
- Jakobson, R. (1959): On Linguistic Aspects of Translation. In: Browner, Reuben Arthur (Hrsg.): *On Translation*, Harvard University Press, 232–239.

- Keller, L (in prep.): Verständlichkeitsoptimierte Medizinkommunikation - Akzeptabilität und Akzeptanz von Texten in Einfacher und Leichter Sprache. Manuskript der Dissertationsschrift. Universität Hildesheim.
- Kröger, J. (in prep.): Zugänglichkeit von Gesundheitsinformationen. Eine korpusgestützte Analyse am Beispiel des Projekts „medizinische Ratgebertexte in Einfacher Sprache“. Manuskript der Dissertationsschrift. Universität Hildesheim.
- Lang, K. (2021): Auffindbarkeit, Wahrnehmbarkeit, Akzeptabilität. Webseiten von Behörden in Leichter Sprache vor dem Hintergrund der rechtlichen Lage. Berlin: Frank & Timme
- Maaß, C. (2020): Easy Language – Plain Language – Easy Language Plus. Berlin: Frank & Timme
- Reiß, K./ Vermeer, H.J. (1984): Grundlegung einer allgemeinen Translationstheorie. Berlin et al.: de Gruyter.
- Rink, I. (2019): „Kommunikationsbarrieren.“ In: Maaß, C./Rink, I. (Hrsg.): Handbuch Barrierefreie Kommunikation. Berlin. Frank & Timme
- Rink, I. (2020): Rechtskommunikation und Barrierefreiheit. Zur Übersetzung juristischer Informations- und Interaktionstexte in Leichte Sprache. Berlin: Frank & Timme
- Schubert, K. (2016): „Barriereabbau durch optimierte Kommunikationsmittel: Versuch einer Systematisierung.“ In: MÄLZER, Nathalie (Hg.): Barrierefreie Kommunikation – Perspektiven aus Theorie und Praxis. Berlin: Frank & Timme, 15-33.
- Schulz, R. (in prep.): Medizinkommunikation in Einfacher und Leichter Sprache für Menschen mit kognitiver Beeinträchtigung. Eine empirische Studie zum Verstehen und Behalten medizinischer Fachinformationen. Manuskript der Disserationschrift. Universität Hildesheim.

A study of holistic features in translational language through syntactic dependency networks

Lu Fan

Xi'an Jiaotong University

Complex network approach provides language research with quantitative measures that can capture global features of language. Nevertheless, previous researchers overlooked one type of language, translational language. Translational language has been recognized as a "third code" by some researchers. Among these studies, "translation universals" proposed by Baker (1993) has become the pivotal and the most controversial one. A series of hypotheses on universal features of translation have been put forward ever since. On one hand, some scholars claimed that they have found evidence for hypotheses concerning general linguistic properties of translational language, such as explicitation, simplification and normalization (Baker 1996, Blum-Kulka and Levenston 1983, Vinay and Darbelnet 1958). On the other hand, some researchers argued that these features of translational language are "probabilistic and conditioned" rather than universal (Toury 2004). Because most researchers focused on individual language structures, connectives (Marco 2018), for example, instead of overall features, such evidence is inconclusive. Thus, the independence of translational language as a third code and its universal features still need to be validated. To study features of translational language from a holistic point of view, Fan and Jiang introduced dependency distance and direction into translation studies (Fan and Jiang 2019). Their study proved the independence of translational language through dependency analysis, and dependency analysis has been shown to be applicable for translation studies. As syntactic dependency network provides translation studies with another holistic perspective in addition to dependency analysis, it is of much necessity and significance to probe into features of translational language through this approach.

In this study, we intend to examine this independence and explore comprehensively the features of translational language. We aim to address the following three research questions: (1) Do small-world and scale-free properties appear in syntactic network of translational language? (2) Is there any difference between translational language and its source language in network parameters? (3) Is there any difference between translational language and native language in network parameters? To validate the independence of translational language and comprehensively explore the universal features of it including S-type plus T-type features, we built a corpus with comparable and parallel language materials in it. The language pair selected is that of Chinese and English, from the Sino-Tibetan and Indo-European language families, respectively. To be specific, the corpus contains the original Chinese *Report on the Work of the Government* (RWG) delivered at an annual China National People's Congress by premiers of China's state council and its English translation (RWG-en), American presidential *State of the Union* (SOU) and its Chinese translation (SOU-ch). A total of 16 texts from the same time span (2014-2017) were selected, with 67,620 English word tokens and 117,396 Chinese characters, respectively.

In the syntactic dependency networks of this study, the nodes are words and the edges are dependency relations between words. A dependency-annotated treebank can be converted into a syntactic network. First, we transformed the corpus into dependency-annotated treebanks. The texts were automatically parsed by the Stanford Parser (3.6.0), a natural language processing program developed by Stanford University. We then manually checked the parsed results to make sure the texts were parsed following the same annotation scheme. To construct syntactic dependency networks, we programmed the output of dependency tagging into an EXCEL format, and then converted it into dependency networks by a network converting software Createpajek.

Among plenty of network properties, the most frequently observed ones are the number of nodes and edges, average degree, average path length, clustering coefficient and non-trivial statistical patterns (scale-free and small-world properties). To verify the small-world property of translational language, we need to compare the original network with its counterpart random network. We adopted Erdős-Rényi random network, which has the same number of nodes and edges as the original one, but links among the nodes are random and all the nodes have the same probability to be linked. A network is a small-world network if it shows almost as small average path length (L) as and far greater clustering coefficient (C) than its random network, i.e. $L \approx L_{\text{random}}$, $C > C_{\text{random}}$ (Watts and Strogatz 1998). A small-world network presents a low degree of separation between nodes and high level of clustering. This property facilitates communication between nodes and thus facilitates mental navigation, if the language network can be viewed as a model of the mental representation of linguistic knowledge (Cong and Liu 2014).

We investigated macroscopically translational language from English into Chinese and from Chinese into English by comparing with its source language and native language through syntactic dependency networks. Statistical methods like non-parametric Mann-Whitney test were adopted to compare the data. The results show that: (1) translational language presents small-world and scale-free properties like most languages do; (2) however, it is independent of and different from both source language and native language in terms of other network parameters; (3) its specific network parameters, i.e. the number of nodes and edges, average degree $\langle k \rangle$, clustering coefficient (C) and average path length (L), show values eclectic between source language and native language, and this eclectic tendency may be regarded as a new candidate for universal features of translational language, which certainly needs further validation in other genres and language pairs. This study also corroborates that quantitative linguistic method of complex network approach can be well utilized in the study of translational language.

References

- Abramov, O. and Mehler, A. (2011) Automatic language classification by means of syntactic dependency networks. *Journal of Quantitative Linguistics*, 18(4), pp. 291-336.
- Baker, M. (1993) Corpus linguistics and translation studies: implications and

- applications. in Baker, M., Francis, G. and Tognini-Bonelli, E., (eds.) *Text and Technology: In Honour of John Sinclair*, Amsterdam and Philadelphia: John Benjamins. pp. 233-250.
- Baker, M. (1996) Corpus-based translation studies: the challenges that lie ahead. in Somers, H., (ed.) *Terminology, LSP and Translation: studies in language engineering: in honour of Juan C. Sager*, Amsterdam and Philadelphia: John Benjamins. pp. 175-186.
- Barabasi, A. L. and Albert, R. (1999) Emergence of scaling in random networks. *Science*, 286, pp. 509-12.
- Baronchelli, A., Ferrer-i-Cancho, R., Pastor-satorras, R., Chater, N. and Christiansen, M. H. (2013) Networks in Cognitive Science. *Trends in Cognitive Sciences*, 17(7), pp. 348-360.
- Beckner, C., Blythe, R., Bybee, J. L., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D. and Schoenemann, T. (2010) Language is a complex adaptive system: position paper. *Language Learning*, 59(s1), pp. 1-26.
- Blum-Kulka, S. and Levenston, E. (1983) Universals of lexical simplification. in Faerch, C. and Casper, G., (eds.) *Strategies in Inter-language Communication*, London and New York: Longman. pp. 119-139.
- Chen, X. and Liu, H. (2016) Function Nodes in Chinese Syntactic Networks. *Understanding Complex Systems*, 99, pp. 187-201.
- Chesterman, A. (2004) Hypotheses about translation universals.
- Cong, J. and Liu, H. (2014) Approaching human language with complex networks. *Physics of Life Reviews*, 11(4), pp. 598-618.
- Duff, A. (1981) *The Third Language : recurrent problems of translation into English : it ain't what you do, it's the way you do it*, Oxford: Pergamon Press.
- Fan, L. and Jiang, Y. (2019) Can dependency distance and direction be used to differentiate translational language from native language? *Lingua*, 224, pp. 51-59.
- Ferrer-i-Cancho, R. (2005) The structure of syntactic dependency networks: insights from recent advances in network theory. in Altmann, G., Levickij, V. and Perebyinis, V., (eds.) *Problems of Quantitative Linguistics*, Chernivtsi: Ruta. pp. 60-75.
- Ferrer-i-Cancho, R., Solé, R. V. and Köhler, R. (2004) Patterns in syntactic dependency networks. *Physical Review E*, 69(1), pp. 051915.
- Frawley, W. (1984) Prolegomenon to a theory of translation. in Frawley, W., (ed.) *Translation: literary, linguistic, and philosophical perspectives*, London: Associated University Presses. pp. 159-175.
- Han, H. and Jiang, Y. (2017) Rethinking translation in the light of complex adaptive system theory. *Chinese Translators Journal*, 38(02), pp. 19-24+128.
- House, J. (2008) Beyond intervention: universals in translation? *trans-kom*, 1(1), pp. 6-19.
- Hudson, R. (2007) *Language Networks: The New Word Grammar*, Oxford: Oxford University Press.
- Hudson, R. (2010) *An Introduction to Word Grammar*, Cambridge: Cambridge University Press.
- Jiang, J., Yu, W. and Liu, H. (2019) Does scale-free syntactic network emerge in

- second language learning? *Frontiers in Psychology*, 10, pp. 925.
- Jin, H. and Liu, H. (2016) Chinese Writing of Deaf or Hard-of-Hearing Students and Normal-Hearing Peers from Complex Network Approach. *Frontiers in Psychology*, 7.
- Lamb, S. M. (1999) *Pathway of the Brain: the Neurocognitive basis of language*, Amsterdam, Philadelphia: John Benjamins.
- Larsen-Freeman, D. and Lynne, C. (2008) *Complex Systems and Applied Linguistics*, Oxford: Oxford University Press.
- Liu, H. (2008) The complexity of Chinese syntactic dependency networks. *Physica A: Statistical Mechanics & Its Applications*, 387(12), pp. 3048-3058.
- Liu, H. (2009) *Dependency Grammar: from theory to practice*, Beijing: Science Press.
- Liu, H. (2010) Translation Studies from Complex Network. *Journal of Beihua University (Social Sciences)*, 11(04), pp. 59-63.
- Liu, H. (2011) Linguistic Networks: Metaphor or Tool? *Journal of Zhejiang University (Humanities and Social Sciences)*, 41(02), pp. 169-180.
- Liu, H. and Cong, J. (2013) Language clustering with word co-occurrence networks based on parallel texts. *Chinese Science Bulletin*, 58(10), pp. 1139-1144.
- Liu, H. and Li, W. (2010) Language clusters based on linguistic complex networks. *Science Bulletin*, 55(30), pp. 3458-3465.
- Liu, H. and Xu, C. (2011) Can syntactic networks indicate morphological complexity of a language? *EPL (Europhysics Letters)*, 93(2), pp. 28005.
- Marco, J. (2018) Connectives as indicators of explicitation in literary translation: A study based on a comparable and parallel corpus. *Target*, 30(3), pp. 87-111.
- Mauranen, A. (2004) Corpora, universals and interference. in Mauranen, A. and Kujamäki, P., (eds.) *Translation Universals : Do They Exist?*, Amsterdam and Philadelphia: John Benjamins. pp. 65-82.
- Motter, A. E., de Moura, A. P. S., Lai, Y.-C. and Dasgupta, P. (2002) Topology of the conceptual network of language. *Physical Review E*, 65(6), pp. 065102.
- Newman, M. E. J. (2003) The structure and function of complex networks. *Siam Review*, 45(2), pp. 167-256.
- Risku, H., Windhager, F. and Apfelthaler, M. (2013) A dynamic network model of translatorial cognition and action. *Translation Spaces*, 2(2), pp. 151-182(32).
- Saussure, F. D. (1959) *Course in General Linguistics*, New York: Philosophical Library.
- Solé, R. V., Corominas-Murtra, B., Valverde, S. and Steels, L. (2010) Language networks: Their structure, function, and evolution. *Complexity*, 15(6), pp. 20-26.
- Steyvers, M. and Tenenbaum, J. B. (2005) The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, 29(1), pp. 41-78.
- Tesnière, L. (1959) *Éléments de syntaxe structurale*, Paris: Klincksieck.
- Toury, G. (2004) Probabilistic explanations in translation studies: Welcome as they are, would they qualify as universals? in Mauranen, A. and Kujamäki, P., (eds.) *Translation Universals - Do They Exist?*, Amsterdam: John Benjamins. pp. 15-25.
- Trosborg, A. (1997) Translating hybrid political texts. in Trosborg, A., (ed.) *Text*

Typology and Translation, Amsterdam and Philadelphia: John Benjamins. pp. 119-143.

Vinay, J.-P. and Darbelnet, J. (1958) *Comparative Stylistics of French and English. A Methodology for Translation*, Amsterdam: John Benjamins.

Watts, D. J. and Strogatz, S. H. (1998) Collective dynamics of 'small-world' networks. *Nature*, 393, pp. 440-442.

Zipf, G. K. (1949) *Human behavior and the principle of least effort: an introduction to human ecology*, Cambridge, MA: Addison-Wesley Press.

Four explanatory variables in an empirical model of translation

Tatiana Serbina¹, Mario Bisiada², Stella Neumann¹

¹RWTH Aachen, ²Universitat Pompeu Fabra

The present study uses the Geometric Multivariate Analysis (GMA), introduced in Diwersy et al. (2014) and Evert and Neumann (2017), to compare linguistic features of German edited translations not only to the text groups of English originals as well as of comparable German originals, but also to translation manuscripts of the same texts. The study aims at analyzing the interplay of language, register, translation status and editorial intervention visible in the effect of lexico-grammatical features considered in the GMA. Taking into account these dimensions of variation and their role in the weakly supervised multivariate analysis, the paper contributes to the work on an empirical model of translation.

The breadth of work over the recent decades has shown that an empirical model of translation should take into account a variety of different factors, such as the languages concerned, register differences within each language, the socio-cultural norms, the translation status of the text, cognitive processing as well as editorial intervention (e.g. Hansen-Schirra et al., 2017; Carl and Schaeffer, 2017). The present study concentrates on the interaction of the four of these factors. First of all, the lexico-grammatical features of texts are determined by the corresponding language (see contrastive studies, e.g. König and Gast, 2018) and register (Neumann, 2021). The studies on the role of the translation status involve analysis of the translation properties or the so-called translationese (Volansky et al., 2015). These three factors have been also studied in combination in a variety of univariate and multivariate studies (Neumann, 2013; Delaere, 2015; Evert and Neumann, 2017). Finally, the fourth factor is concerned with editorial intervention. A number of recent studies considered editors' influence on translated or non-translated texts drawing attention to the fact that the published versions of texts, typically included in different corpora, do not necessarily contain the linguistic features selected by the authors or translators (e.g. Kruger, 2017; Bisiada, 2019; Serbina et al., 2021). Taking into account the results of these studies, the present paper suggests combining the four variables mentioned above and examining their interaction both in terms of a joint effect of all selected linguistic features as well as the contribution of individual features.

The data set consists of texts from two corpora, namely the Harvard Business Corpus (Bisiada, 2018) and the CroCo Corpus (Hansen-Schirra et al., 2012). The Harvard Business Corpus contains articles published in English in the Harvard Business Review, edited German translations of these articles published in the Harvard Business Manager as well as German translation manuscripts of the same articles. The latter represent translation versions submitted by the translators and forwarded to the publisher by the translation company. The articles could be classified as belonging to the register BUSINESS. The second part of the data set contains English and German originals as well as published German translations taken from the CroCo corpus. The texts belong to the registers of letters to shareholders (SHARE), popular-scientific literature (POPSCI), prepared speeches (SPEECH) and political essays (ESSAY).

English texts from both corpora were POS tagged with the CLAWS 7 tagset using the CLAWS tagger (Garside and Smith, 1997), whereas the German texts were tagged with the STTS tagset (Schiller et al., 1999) using the TreeTagger (Schmid, 1994). GMA is based on the set of lexico-grammatical features introduced in Neumann (2013). These features are automatically extracted and quantified per text using a CQP script (Neumann and Evert, Forthcoming; Fest et al., 2019). The multivariate analysis of texts involves visual inspection and linguistic interpretation of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Preliminary results are based on 187 texts. In accordance with the first steps of GMA, the analysis begins with performance and visual inspection of PCA. Concentrating on the individual variables, the first PCA dimension separates texts according to language, so that English and German texts are grouped on different parts of the axis. Moreover, the variable translation status also plays a role for this dimension: while the English and German originals are grouped on the negative and positive sides of the axis, respectively, the German translations included in the data set are clustered closer to the center. This distribution is in line with the results reported in the study applying GMA to several registers of the CroCo corpus: Evert and Neumann (2017) attribute this distribution to the shining-through effect of translations. The variation along the second dimension could be linked to the explanatory variable of register. As noted by Serbina et al. (2021), BUSINESS and SHARE registers are separated most distinctly. While the third dimension does not appear to be very informative, the variable of editorial intervention may play a minor role in the fourth PCA dimension.

In the next step, we perform the supervised LDA using translation status as a discriminant. Since the discriminant has two levels, the resulting model contains only one LDA dimension. This dimension separates texts into originals and translations. However, the separation is not very clear, since German translations from the CroCo corpus appear to be more similar to the German originals than both versions of translations from the Harvard Business Corpus. Considering the results of the PCA, it might be meaningful to add language as the discriminant to this analysis.

Taking into account the combined effect of lexico-grammatical features, these initial results suggest that such factors as translation status, language and register appear to have the most profound effect on the text distribution, whereas editorial intervention plays a more marginal role in this holistic perspective. To further study the role of this particular set of explanatory variables, the next step of the analysis will consider in more detail the distribution of those features that contribute most to different dimensions of variation reported above.

References

- Mario Bisiada. 2018. Editing nominalisations in English–German translation. *The Translator*, 24(1):35–49.
- Mario Bisiada. 2019. Translated language or edited language? A study of passive constructions in translation manuscripts and their published versions. *Across Languages and Cultures*, 20(1):35–56.
- Michael Carl and Moritz Schaeffer. 2017. Sketch of a noisy channel model for the translation process. In Silvia Hansen-Schirra, Oliver Czulo, and Sascha

- Hofmann, editors, *Empirical modelling of translation and interpreting*, pages 71–116. Language Science Press, Berlin.
- Isabelle Delaere. 2015. *Do Translations Walk the Line? Visually Exploring Translated and Non-Translated Texts in Search of Norm Conformity*. Ph.D. thesis, Ghent.
- Sascha Diwersy, Stefan Evert, and Stella Neumann. 2014. A weakly supervised multivariate approach to the study of language variation. In Benedikt Szmrecsanyi and Bernhard Wälchli, editors, *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech*, pages 174–204. De Gruyter, Berlin.
- Stefan Evert and Stella Neumann. 2017. The impact of translation direction on characteristics of translated texts: A multivariate analysis for English and German. In Gert de Sutter, Marie-Aude Lefer, and Isabelle Delaere, editors, *Empirical Translation Studies: New Theoretical and Methodological Traditions*, pages 47–80. Mouton de Gruyter, Berlin.
- Jennifer Fest, Arndt Heilmann, Oliver Hohlfeld, Stella Neumann, Helge Reelfs, Marco Schmitt, and Alina Vogelgesang. 2019. Determining response-generating contexts on microblogging platforms. In *Proceedings of the 15th Conference on Natural Language Processing (KONVENS)*, pages 171–182.
- Roger Garside and Nicholas Smith. 1997. A hybrid grammatical tagger: Claws4. In Roger Garside, Geoffrey Leech, and Anthony McEnery, editors, *Corpus Annotation: Linguistic information from computer text corpora*, pages 102–121. Longman, London.
- Silvia Hansen-Schirra, Stella Neumann, and Erich Steiner. 2012. *Cross-linguistic Corpora for the Study of Translations: Insights from the Language Pair English- German*. de Gruyter, Berlin.
- Silvia Hansen-Schirra, Jean Nitzke, and Katharina Oster. 2017. Predicting cognate translation. In Silvia Hansen-Schirra, Oliver Czulo, and Sascha Hofmann, editors, *Empirical modelling of translation and interpreting*, pages 3–22. Language Science Press, Berlin.
- Haidee Kruger. 2017. The effects of editorial intervention: Implications for studies of the features of translated language. In Gert De Sutter, Marie-Aude Lefer, and Isabelle Delaere, editors, *Empirical Translation Studies: New Methodological and Theoretical Traditions*, pages 113–156. de Gruyter, Berlin.
- Ekkehard König and Volker Gast. 2018. *Understanding English-German Contrasts*, 4 edition. Erich Schmidt Verlag, Berlin.
- Stella Neumann. 2013. *Contrastive Register Variation: A Quantitative Approach to the Comparison of English and German*. de Gruyter, Berlin.
- Stella Neumann. 2021. Register and translation. In Mira Kim, Jeremy Munday, Zhenhua Wang, and Pin Wang, editors, *Systemic functional linguistics and translation studies*, pages 65–82. Bloomsbury, London.
- Stella Neumann and Stefan Evert. Forthcoming. A register variation perspective on varieties of english. In Elena Seoane and Douglas Biber, editors, *Corpus based approaches to register variation*. de Gruyter, Berlin.
- Anne Schiller, Simone Teufel, Christine Stöckert, and Christine Thielen. 1999. *Guidelines für das Tagging Deutscher Textcorpora mit STTS*. Universität Stuttgart, Universität Stuttgart.

- Helmut Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In Proceedings of the International Conference on New Methods in Language Processing.
- Tatiana Serbina, Mario Bisiada, and Stella Neumann. 2021. Linguistic profiles of translation manuscripts and edited translations. In Proceedings for the First Workshop on Modelling Translation: Translatology in the Digital Age, pages 34–45. Association for Computational Linguistics.
- Vered Volansky, Noam Ordan, and Shuly Wintner. 2015. On the features of translationese. *Digital Scholarship in the Humanities*, 30(1):98–118.

Unique Item Hypothesis Revisited from the Perspective of Translation Directionality

Juhong Zhan, Yue Jiang, Baicen Jin, Tianyuan Yang
Xi'an Jiaotong University

Unique item hypothesis (UIH) is one of the translation universal tendencies proposed in corpus-translation studies. It claims that unique items in the target language will tend to be underrepresented (or less frequently used) in the translated texts when compared with those in the comparable non-translated texts.

Translation directionality concerns whether translators are working from a foreign language into their mother language or vice versa. Translation done from a foreign language into one's mother tongue is called native translation, otherwise non-native translation. The present study aims to revisit UIH from the perspective of translation directionality in Chinese-English(C-E) language pair.

Unique items are defined as target-language-specific items that lack straightforward translation counterparts or equivalents in the source language. In C-E translation, **phrasal verbs** (PV) possess all the characteristics of unique items loosely defined by Andrew Chesterman (2004). Morphologically, phrasal verb is not hard to define. It is a structure consisting of a lexical verb and an adverbial particle that function as a single unit both lexically and syntactically (Darwin & Gray 1999). Syntactically, the adverbial particle and the lexical verb making up the PVs can be either contiguous, like *go over*, or noncontiguous, like *give sth./sb. Away* (Gardner & Davies 2007). Semantically, PVs are classified into three categories with semantic transparency ranging from the least to the most: directional (literal or transparent) like *stand up* and *take away*, aspectual (completive) like *burn down* and *eat up*, and idiomatic (figurative, opaque) ones like *face off* and *figure out* (Riguel 2014). PVs are peculiar of English and have very few parallels in Chinese, and thus are regarded as unique items in C-E translations. Therefore, we'll use PV as the language feature to investigate whether UIH holds truth in C-E translations and whether translation directionality plays a role in the representation of unique items. Specifically, we aim to answer following research questions: 1) Are PVs underrepresented in C-E translated texts when compared with English non-translated texts? 2) Is there a significant difference in the use of PVs between different translation directions? 3) If the answer to the second question is yes, what are the differences?

We built a parallel corpus of Lu Xun's short stories and their English translations done by two native and two non-native translators, and a comparable reference corpus of BNC short stories as the non-translated reference.

Lu Xun (1881-1936) was a pioneer and founder of modern Chinese literature, a short story writer, essayist, poet, and literary critic. He has been described by Japanese Nobel laureate Kenzaburō Ōe as "the greatest writer Asia produced in the twentieth century."(Kenzaburō Ōe, 1995) As an intellectual titan, a literary colossus and an ideological iconoclast, "Lu Xun remains the most translated and studied" of modern Chinese writers (Wang, 2011).

The present study chose Lu Xun's short stories as the source texts for two important concerns, his distinctive language style and the abundance of translations

of his works. In Chinese literary history, Lu Xun was the first to write in vernacular Chinese, the modern colloquial and everyday language used by ordinary people, using large quantity of idioms and colloquialisms (Wang, 2011). Lu Xun's dialectal tone and colloquial style, featured by high level of colloquialness and idiomaticity, is hard to convey through translation. As phrasal verb is the typical feature of colloquial language and daily English, English translations of Lu Xun's works are expected to contain many PVs, if they are to preserve the colloquiality of the source language.

As to the translators, both Wang and Yang are very experienced and proficient bilingual translators. They received higher education in English-speaking countries and were exposed to English language environment for years. Their translation can represent the highest level of Chinese to English non-native translations. Likewise, both the two native translators William Lyell and Julia Lovell enjoy very high fame in translating Chinese literature into English. Their translations are believed to symbolize a highest quality of C-E translational language.

We designed and built a parallel corpus composed of 10 Chinese short stories and their correspondent English translations done by native and non-native translators. It contains a sub-corpus of source texts including 84018 Chinese words and four sub-corpora of target texts including 261888 English words in total. The target texts are cleaned, tagged with CLAWS7 and aligned at the sentence level with the cleaned source texts. To build a parallel non-translated corpus as reference, we randomly sampled 40 BNC short stories from BNC-fiction, totaling 371,155 words. BNC short stories are also cleaned and tagged with CLAWS7.

Combining quantitative and qualitative analysis, the research follows these steps:

1. Counting the occurrences of phrasal verbs by searching “_RP”

Firstly, we searched with the *concordance* function of AntConc3.5.7w for phrasal verbs in each text. In CLAWS7 tagset, the adverbial preposition in phrasal verbs is tagged as RP, which goes with a verb to form a phrasal verb. We noticed that the tagging system is not 100% accurate, and sometimes RP just doesn't go with a verb in cases like 'all the way *down* to the floor', in which *down* is an adverbial particle but not a part of phrasal verbs. After searching and counting _RP in each translational text by different translators, we manually screen out the unwanted concordances.

2. Performing statistical significance testing

After counting the occurrences of PV in each text, we conducted T-test using SPSS24.0 to test whether there is a significant difference in the frequencies of phrasal verbs between translated texts & non-translated texts, and between native & non-native translations.

3. Programming to generate the rank-frequency lists of PVs

Then we programmed with python to extract all the phrasal verbs, lemmatize them and generate PV lists. PVs under CLAWS tagging system are annotated as VV0, VVD, VVG (including VVGK), VVI, VVN and VVZ as lexical verbs followed by RP standing for adverbial prepositions. The number of intervening words between the lexical verb and the adverbial particle is set between 0-6 (Gardner & Davies, 2007).

4. Performing semantic analysis of PV key words

As keyword analysis (just like key word) can reflect which word is overused or underused in different corpus. We used the *keyword* function of Wordsmith tools 7.0

to generate key PV lists of the native and non-native corpus by using the PV lists generated in step 3 and made some semantic analysis qualitatively.

Table 1 Results of the T-test significance analysis

Texts	BNC	Translations (vs. BNC)	Native (vs. BNC)	Non-native (vs. BNC)	Native vs. Non-native
N	40	40	20	20	
Mean	12.0	15.8	19.2	12.3	
SD	5.1	4.6	3.6	2.3	
<i>P</i>		.001***	.000***	.656	.000***

Table 2 Significance analysis within and between translation directions

	Translator	N	Mean	Sig (2-tailed)
Within group	Lovell	10	18.9	.701
	Lyell	10	19.5	
Within group	Wang	10	11.8	.281
	Yang	10	12.9	
Between groups	Native	20	19.2	.000***
	Non-native	20	12.4	

1. Frequency of PV in BNC short stories is averaged 12.0%, with the peak point at 25.4%, and the valley point at 3.5%. Frequency of PV in Luxun translations is averaged 15.8%, with the peak at 23.3% and the valley at 9.1%. And the standard deviations in BNC is 5.1, higher than 4.6 in translations, suggesting that PV use across BNC short stories is more spread out, and the data is more dispersed than translations(see table 1).

2. Frequency of PV in native translations is averaged 19.2%, with the peak point at 23.3%, and the valley point at 13.4%. Frequency of PV in non-native translations is averaged 12.3%, with the peak at 16.5% and the valley at 9.1%. And the standard deviations in native translations is 3.6, higher than 2.3 in non-native translations, suggesting that PV use in native translations is more spread out, and the data is more dispersed than non-natives(see table 1).

3. The independent Samples t-test shows there is significant difference in the frequency of PVs between English translations of Luxun's short stories and BNC short stories ($P=.001 < .05$); and there is significant difference in the use of PVs between native translations and BNC short stories($P=.000 < .05$); but there is no significant difference in the use of PVs between non-native translations and BNC fictions($P=.656 > .05$) (see table 2);

4. There is significant difference in the use of PVs between native and non-native translations ($P=.000 <.05$); within the same translation direction, no significant difference is found in the use of phrasal verbs ($P>.05$); between translation directions, a significant difference is found ($P <.05$). (see table 2)

5. Semantic analysis of top 10 key PVs reveals non-natives prefer semantically semi-transparent and transparent phrasal verbs. The key PVs used by non-native translators are composed of a dynamic action verb like *go, come, rush, walk, and look*, plus a commonly used particle. Among them, *stand up, go out, rush out, go in, come in, look up, and come back* are directional, and *walk on* and *get back* are aspectual. They are comparatively simpler in form and explicit and transparent in meaning. Among the ten key PVs used by native translators, only *come along* and *stare up* are aspectual, all the others including *work up, take back, figure out, head back, let out, scrub up, head off* and *throw in* are idiomatic. Idiomatic PVs are non-compositional and semantically opaque. Their meanings are more implicit and figurative, unrelated to the component verb or particle. Therefore, we tend to propose that non-native translators prefer literal PVs than idiomatic ones, while native translators' use of different categories of PVs is more balanced, as the keyness and frequency of their key PVs are both much lower than that of non-native translators. non-native translators, just like non-native English language users, use fewer idiomatic PVs than non-native translators.

Based on the results above, our answers to the research questions are: 1) PVs are significantly over-represented in C-E translated texts when compared with English non-translated texts, and this overrepresentation is mainly attributed to the remarkable use of PVs by native translators; 2) A significant difference is found in the use of PVs by translators of different directionality, while no significant difference is found within the same direction. 3) Use of phrasal verbs diverges between C-E native and non-native translations, specifically in their frequency and the degree of semantic transparency.

The above findings contradict with the unique item hypothesis in that phrasal verbs are slightly overrepresented in C-E translations than English original, which is just the opposite to what is hypothesized. But when translation direction is considered, we find that the representation of PVs diverges between different directions, that is, PVs are slightly underrepresented in non-native translations but remarkably overrepresented in native translations. This result has falsified unique item hypothesis in general, or at least, we find that unique item hypothesis seems constrained by translation direction.

We're seeking to interpret these differences from the perspective of contrastive cognitive processes of native and non-native translators.

It is hard to find more translators of different directionality who translate the same source texts. Thus, the results of the present study need to be verified in a bigger corpus and among more translators in the future.

References

- Chesterman A. Beyond the particular[J]. Translation universals: Do they exist, 2004, 33: 49.
- Darwin C M, Gray L S. Going after the phrasal verb: An alternative approach to classification[J]. Tesol Quarterly, 1999, 33(1): 65-83.

- Gardner D, Davies M. Pointing out frequent phrasal verbs: A corpus-based analysis[J]. TESOL quarterly, 2007, 41(2): 339-359.
- Riguel E. Phrasal verbs, "the scourge of the learner"[C]//9th Lancaster University Postgraduate Conference in Linguistics & Language Teaching. 2014, 9: 1-20.
- Wang, Baorong, and M. Phil. 2011. Lu Xun's fiction in English translation: the early years. PhD diss. University of Hong Kong.
- Gardner D, Davies M. Pointing out frequent phrasal verbs: A corpus-based analysis[J]. TESOL quarterly, 2007, 41(2): 339-359.

Activity as a discriminant of direct and inverse translations: A corpus-based study of eight translations of *Li Sao*

Zuohao Xu and Yue Jiang
Xi'an Jiaotong University

Direct translations (DTs) and inverse translations (ITs) represent L1 and L2 translators' group styles, respectively (Marmaridou 1996; Lonsdale 2004). However, it remains controversial whether there is an explicit stylistic discriminant of them. The dynamic and synergic nature of texts was little noticed by most prior studies but is valued by quantitative linguistics (Köhler 2008; Liu 2017). Activity, as a stylistic parameter of quantitative linguistics, refers to the ratio of verbs to the sum number of verbs and adjectives (Zörnig *et al.* 2015: 4; Liu 2017: 136). Albeit that adjectives are dispensable modifiers parasitic on nouns (Kostić and Katz 1987), verbs are the essential components of sentences (Baker 2003) because they constitute the cores of argument structures (Goldberg 1995) and can restrict forthcoming themes (Altmann & Kamide 1999). Therefore, humans tend to prioritize verbs over adjectives in language processing, esp. under extreme cognitive pressure (Jia & Liang, 2020). Based on this observation, we assume that translators tend to produce a text with higher activity value if they are under more cognitive pressure or have less cognitive resources available. Since activity not only reflects the dynamic and synergic nature of verbs and adjectives but also has cognitive significance, the present study adopts it as a stylistic parameter to explore the difference between DTs and ITs, and that between them and the original English.

The corpus used in this study consists of eight English translations of *Li Sao* poem by Qu Yuan, four of which are DTs and four are ITs. The DTs were translated by James Legge, David Hawkes, Stephen Owen and Gopal Sukhu, each with 3991, 3469, 3312 and 3398 tokens, respectively, and ITs were translated by Xianyi Yang, Yuanchong Xu, Dayu Sun and Zhenying Zhuo, with 2898, 2982, 3399, 3378 tokens, respectively. BNC's poetry sub-corpus is used as a reference corpus (255,630 tokens). Three research questions are to be addressed: 1) Can activity help differentiate DTs from ITs? 2) How do the two differ from each other in terms of activity? 3) What are the differences between DTs/ ITs and the English original poetry in terms of activity?

Two-step Cluster Analysis was performed to verify whether the eight translations can be properly clustered into two groups in terms of activity, namely, DTs and ITs as they are supposed to belong to. Then a Mann-Whitney Test was performed to determine whether the difference between DTs and ITs is significant. Finally, the difference between DTs/ ITs and the English original poetry is visualized with a bar graph. The results show that 1) the eight translations are properly clustered into the supposed groups of DTs and ITs (Average Silhouette= 0.6); 2) DTs' activity (M= 0.77, SD= 0.019) is significantly higher than ITs' (M= 0.71, SD= 0.013) (Asymp. Sig. (2-tailed)= 0.021, Exact Sig. [2*(1-tailed Sig.)]=0.029); 3) DTs' activity is higher than that of the original English (0.735289) while ITs' activity is lower than that of the original English.

Result 1) suggests that activity can indeed be used to stylistically differentiate DTs and ITs. Result 2) indicates that, compared with the inverse translators, the direct translators have less cognitive resources when producing the target texts, which compels them to allocate more cognitive efforts to verbs, the more important word class. If translation process can be sub-divided into two stages like reading the source text and producing the target text, the direct translators tend to invest more cognitive efforts in the reading and decoding of the source text because they have less prior knowledge of it. On the other hand, given that the total human cognitive capacity is limited, when more cognitive efforts are invested in the source text, less cognitive efforts will accordingly be left to produce the target text. Result 3) suggests that the source language shining-through effect (Teich 2003: 207) is more prominent in DTs while the normalization (Baker 1996: 183) is more explicit in ITs. In C-E language pair, higher activity ($M= 0.842$, $SD=0.013$) is a prominent feature of Chinese while lower activity ($M= 0.692$, $SD= 0.057$) is a prominent feature of English (Xu & Jiang 2021). DTs' activity exhibits a close affinity to that of the Chinese source language, which may plausibly suggest that source language shining-through effect plays a dominant role in DTs. In contrast, ITs' activity exaggerates the lower activity of English, the target language, plausibly indicative of a normalizing tendency. Prior studies report that direct translators, viz. L1 translators, are more concerned with source text aspects such as style or vocabulary (Ferreira *et al.* 2018: 112) whereas L2 readers retain more surface linguistic information of the reading materials and rely on it more than L1 readers (Bordag *et al.* 2021: 9). When reading the source text, as direct translators are L2 readers, they retain more surface linguistic information (such as activity) of the source text and rely more on it than L1 readers, which induces the source language shining-through effect. So, in essence, source language shining-through effect is an effort-sparing way to reduce the translator's cognitive pressure, esp. salient in DTs, because direct translators have less prior knowledge of the source text and thus have to rely more on its surface linguistic information. This is in line with the Principle of Least Effort (Zipf 1949). Unlike direct translators, inverse translators, having more prior knowledge of the source language system, can process the source text easily and thus have little need to rely on its surface linguistic information. However, less familiar with the target language system, inverse translators have to abide by its norms too closely for risk aversion (Pym 2015), which may result in normalization. These findings reveal the differences between DTs and ITs in cognitive pressure and risk aversion strategy, and may hopefully shed some light on the study of language contact and the modeling of automatic identification of L1 and L2 translators' group styles.

References

- Altmann, G.T.M. & Y. Kamide. 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73 (3): 247-264.
- Baker, M. 1996. Corpus-based translation studies: The challenges that lie ahead [A]. In H. Somers (ed.). *Terminology, LSP and Translation* [C]. Amsterdam: John Benjamins. 175-186.

- Baker, M. C. 2003. *Lexical Categories: Verbs, Nouns and Adjectives* [M]. Cambridge: Cambridge University Press.
- Bordag D, Opitz A, Polter M, Meng M. 2021. Non-native Readers Are More Sensitive to Changes in Surface Linguistic Information than Native Readers [J]. *Bilingualism: Language and Cognition* 24 (4): 1–13.
- Ferreira, A., A. Gottardo & J. W. Schwieter. 2018. Decision-making processes in direct and inverse translation through retrospective protocols [J]. *Translation, Cognition & Behavior* 1 (1): 98–118.
- Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure* [M]. Chicago: University of Chicago Press.
- Jia, H. & J. Liang. 2020. Lexical category bias across interpreting types: Implications for synergy between cognitive constraints and language representations [J]. *Lingua* 239: 1-13.
- Köhler, R. 2008. Synergetic linguistics [A]. In R. Köhler, G. Altmann and R. G. Piotrowski (eds.) *Quantitative Linguistik / Quantitative Linguistics. Ein internationales Handbuch / An International Handbook*. 760 – 774. Berlin, Boston: De Gruyter Mouton.
- Kostić, A. & L. Katz. 1987. Processing differences between nouns, adjectives, and verbs [J]. *Psychological Research* 49 (4): 229-236.
- Liu, H. T. 2017. *An Introduction to Quantitative Linguistics* [M]. Beijing: The Commercial Press. [刘海涛, 2017, 计量语言学导论[M].北京: 商务印书馆].
- Lonsdale, A. B. 2004. Direction of translation: directionality [A]. In M. Baker (ed.) *Routledge Encyclopedia of Translation Studies* [C]. Shanghai: Shanghai Foreign Language Education Press. 63-67.
- Marmaridou, A. S. 1996. Directionality in translation processes and practices [J]. *Target* 8 (1): 49-73.
- Pym, A. 2015. Translating as risk management [J]. *Journal of Pragmatics* 85: 67-80.
- Teich, E. 2003. *Cross-linguistic Variation in System and Text: A Methodology for the Investigation of Translations and Comparable Texts* [M]. Berlin: Mouton de Gruyter.
- Xu, Z. H. & Y. Jiang. 2021. *Activity of translational Chinese: A study based on three online corpora* [J]. *Foreign Language Teaching and Research* 53 (1): 113-123. [徐佐浩、蒋跃, 2021, 翻译汉语的活动度——基于在线语料库的研究[J], 《外语教学与研究》53 (1): 113-123.]
- Zörnig, P., K. Stachowski, I. Popescu, T. Miyangah, P. Mohanty, E. Kelih, R. Chen & G. Altmann. 2015. *Descriptiveness, Activity and Nominality in Formalized Text Sequences* [M]. Lüdenscheid: RAM-Verlag.
- Zipf, G. K. 1949. *Human Behavior and the Principle of Least Effort* [M]. Cambridge: Addison-Wesley Press.

Cognitive Effort Variation Between Translators of Language Pairs of Different Scripts and Text Directions

Ali Saeedi; Longhui Zou
Kent State University

Language pairs written with different scripts entail dissimilar cognitive efforts during the translation process. An increased effort for these languages is to be expected compared to the case of translation between the same-script language pairs (Lacruz et al., 2016). Also, many studies showed effect of native language direction on spatial bias in reading comprehension (Barrett et al., 2002; Boroditsky, 2001; Chatterjee et al., 1999; Spalek & Hammad, 2005). In general, text direction and language orientation create a spatial bias where right-oriented readers demonstrate right bias and left-oriented readers demonstrate left bias (Mashat, 2017). Much research has focused on the relation between different themes of language specificity and translation difficulty (Campbell, 1999; Gile, 2005, 2008; Wang & Zou, 2018). Nonetheless, little research has centered on the relation of orthography and text direction with translation difficulty. Even more scarce research has compared different language pairs in regard to the orthography and text direction, for instance, English into Spanish, English into Chinese, and English into Arabic. Drawing on these points, this study aims at exploring whether translating the source text (ST) and target text (TT) of different scripts will involve more cognitive effort than when the translator's language pair scripts are similar. Another aim of this study is to investigate whether the difference in text direction between the language pair is another trigger of more cognitive effort in translation. Showing that languages of different scripts and text directions impose an added challenge to translators of these language pairs can justify attempts to finding special strategies and techniques of translation for such language pairs. This will, in turn, open the room for researching and introducing pedagogical approaches to translator training that might align with our findings.

In this study, we compare three datasets available at the Center for Research and Innovation in Translation and Translation Technology's Translation Process Research Database (CRITT TPR-DB) (see Carl et al., 2016). The three datasets are BML12 (English to Spanish), RUC17 (English to Chinese) and AR19 (English to Arabic). These three datasets share an identical set of English STs, which allow us to control the linguistic complexity, script variation, and text direction of the STs among all datasets. The STs include six English texts from Encyclopedia (four news articles and two sociology articles). The total number of words for the six STs is 847 and the total number of segments is 41. BML12 dataset involves 32 professional translators of English into Spanish. This dataset is used in this study as the control group that the other groups will be compared to first and foremost since Spanish has the same script and text direction of English. RUC17 dataset involves 21 professional translators of English into Chinese. This dataset represents the translation process of the translators working with a target language of different script but similar text direction compared to English. AR19 dataset comprises 15 professional translators of English into Arabic. This dataset denotes the translation process of the translators working with the target language of different script and

different text direction from English. All translators in these three datasets have English as their second language of command and are native speakers of their respective target language. The behavioral data of all participants in these three datasets including eye tracking and key logging data were retrieved from CRITT TPR-DB.

We operationalize cognitive effort among these three selected datasets by commonly used indicators such as the translation duration, the total number of fixations, the total duration of fixations, the duration of individual fixations, the transitions across ST and TT during translation, etc. We first compare the behavioral data of translators of English-into-Spanish language pair with those of translators of English-into-Chinese and English-into-Arabic language pairs for possible indicators of different cognitive effort with language pairs of different scripts. For the translators of English into Spanish, our data indicate an overall tendency of less cognitive effort compared to the other two language pairs. For instance, the translation duration is relatively shorter and the total fixation durations are fairly shorter. These findings can be attributed to the similarity of script of Spanish and English. Subsequently, we compare English-into-Chinese language pair with English-into-Arabic language pair for possible markers of different cognitive effort with language pairs of different text directions. The translators of English into Chinese finished the translation task faster than the translators of English into Arabic. They show relatively smaller number of fixations, and faster transitions compared to the translators of English into Arabic. A possible explanation for these findings is that the different text direction of Arabic adds another layer of challenge to translators. We conclude that translation of language pairs of different scripts and/or text direction involves more cognitive effort when compared to language pairs of similar scripts and/or text direction. This higher cognitive effort might indicate greater difficulty in the translation as well.

References

- Barrett, A. M., Kim, M., Crucian, G. P., & Heilman, K. M. (2002). Spatial bias: Effects of early reading direction on Korean subjects. *Neuropsychologia*, *40*(7), 1003-1012.
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive psychology*, *43*(1), 1-22.
- Campbell, S. (1999). A cognitive approach to source text difficulty in translation. *Target. International Journal of Translation Studies*, *11*(1), 33-63.
- Carl, M., Bangalore, S., & Schaeffer, M. (2016). New directions in empirical translation process research. *Heidelberg: Springer International Publishing Switzerland. doi, 10, 978-973.*
- Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia*, *37*(4), 395-402.
- Gile, D. (2005). Directionality in conference interpreting: A cognitive view. *Communication & cognition. Monographies*, *38*(1-2), 9-26.
- Gile, D. (2008). Issues in research into conference interpreting. In *Übersetzung* (pp. 767-779). De Gruyter Mouton.

- Lacruz, I., Carl, M., Yamada, M., & Aizawa, A. (2016). Pause metrics and machine translation utility. *Proceedings of the NLP*, 1213-1216.
- Mashat, A. A. (2017). *An Eye Tracking Study to Investigate the Influence of Language and Text Direction on Multimedia*. Old Dominion University.
- Spalek, T. M., & Hammad, S. (2005). The left-to-right bias in inhibition of return is due to the direction of reading. *Psychological Science*, 16(1), 15-18.
- Wang, B., & Zou, B. (2018). Exploring language specificity as a variable in Chinese-English interpreting. A corpus-based investigation. In *Making way in corpus-based interpreting studies* (pp. 65-82). Springer.